

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: July 31, 2003, 11:02:37 ; Search time 2765 Seconds  
(without alignments)  
10230.715 Million cell updates/sec  
Title: US-09-992-430B-21  
Perfect score: 972  
Sequence: 1 atgttcaagatacaaaagtc.....ttgagcatctggagatcaat 972  
Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0  
Searched: 2054640 seqs, 14551402878 residues  
Word size : 0  
Total number of hits satisfying chosen parameters: 4109280  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Listing first 1000 summaries

Database : GenEmbl:  
1: gb\_ba:  
2: gb\_htg:  
3: gb\_in:  
4: gb\_om:  
5: gb\_ov:  
6: gb\_pat:  
7: gb\_ph:  
8: gb\_pl:  
9: gb\_pr:  
10: gb\_ro:  
11: gb\_sts:  
12: gb\_sy:  
13: gb\_un:  
14: gb\_vi:  
15: em\_ba:  
16: em\_fun:  
17: em\_hum:  
18: em\_in:  
19: em\_mu:  
20: em\_om:  
21: em\_or:  
22: em\_ov:  
23: em\_pat:  
24: em\_ph:  
25: em\_pl:  
26: em\_ro:  
27: em\_sts:  
28: em\_un:  
29: em\_vi:  
30: em\_htg\_hum:  
31: em\_htg\_inv:  
32: em\_htg\_other:  
33: em\_htg\_mus:  
34: em\_htg\_pln:  
35: em\_htg\_rod:  
36: em\_htg\_mam:  
37: em\_htg\_vrt:  
38: em\_sy:  
39: em\_htgo\_hum:  
40: em\_htgo\_mus:  
41: em\_htgo\_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	23	2.4	151778	2	AC113646	AC113646 Rattus no
C 2	22	2.3	214265	2	AC099573	AC099573 Mus muscu
C 3	21	2.2	62004	2	AC114589	AC114589 Mus muscu
C 4	21	2.2	64969	9	AL355985	AL355985 Human DNA
C 5	21	2.2	150199	9	AC087568	AC087568 Pan trogl
C 6	21	2.2	169574	2	AC111597	AC111597 Rattus no
C 7	21	2.2	177066	9	AC091001	AC091001 Papio cyn
C 8	21	2.2	177893	9	AC005089	AC005089 Homo sapi
C 9	21	2.2	178228	2	AC128082	AC128082 Rattus no
C 10	21	2.2	343550	1	AP003587	AP003587 Nostoc sp
C 11	20	2.1	188	9	HUMLDHX3	M24512 Homo sapien
C 12	20	2.1	811	9	AF401096	AF401096 Homo sapi
C 13	20	2.1	879	9	AF401095	AF401095 Homo sapi
C 14	20	2.1	1047	9	AF401097	AF401097 Homo sapi
C 15	20	2.1	1053	9	AF401094	AF401094 Homo sapi
C 16	20	2.1	1179	9	HUMLDHX	J02938 Human testi
C 17	20	2.1	1204	9	AK058181	AK058181 Homo sapi
C 18	20	2.1	1215	4	VVU19868	U19868 Vulpes vulp
C 19	20	2.1	1243	6	I05686	I05686 Sequence 1
C 20	20	2.1	1250	9	BC019249	BC019249 Homo sapi
C 21	20	2.1	1254	9	HSU13680	U13680 Human lacta
C 22	20	2.1	2745	8	AF257777	AF257777 Vitis vin
C 23	20	2.1	13123	1	AE003872	AE003872 Xylella f
C 24	20	2.1	82625	2	AC097779	AC097779 Rattus no
C 25	20	2.1	87210	8	AB011485	AB011485 Arabidops
C 26	20	2.1	155085	2	AC025620	AC025620 Homo sapi
C 27	20	2.1	157402	2	AC013794	AC013794 Homo sapi
C 28	20	2.1	159998	2	AC074292	AC074292 Homo sapi
C 29	20	2.1	160930	9	AC010093	AC010093 Homo sapi
C 30	20	2.1	165633	2	AC026699	AC026699 Homo sapi
C 31	20	2.1	166973	9	AC084117	AC084117 Homo sapi
C 32	20	2.1	167916	2	AC094048	AC094048 Rattus no
C 33	20	2.1	174151	2	AC011182	AC011182 Homo sapi
C 34	20	2.1	176613	9	AC027708	AC027708 Homo sapi
C 35	20	2.1	177632	2	AC128367	AC128367 Rattus no
C 36	20	2.1	183263	2	AC127896	AC127896 Rattus no
C 37	20	2.1	186000	2	AC131014	AC131014 Rattus no
C 38	20	2.1	189893	9	AC011361	AC011361 Homo sapi
C 39	20	2.1	190352	2	AC097471	AC097471 Homo sapi
C 40	20	2.1	193657	2	AC105552	AC105552 Rattus no
C 41	20	2.1	196410	9	AC027544	AC027544 Homo sapi
C 42	20	2.1	202397	2	AC099232	AC099232 Rattus no
C 43	20	2.1	210533	2	AC099719	AC099719 Mus muscu
C 44	20	2.1	210841	2	AC121840	AC121840 Mus muscu
C 45	20	2.1	220101	2	AC121993	AC121993 Mus muscu
C 46	19	2.0	256	10	AB04820S05	AB047790 Mus muscu
C 47	19	2.0	1621	9	AK058192	AK058192 Homo sapi
C 48	19	2.0	1680	9	AY009108	AY009108 Homo sapi
C 49	19	2.0	1696	9	BC022034	BC022034 Homo sapi
C 50	19	2.0	2221	8	SPISFS1A	M83940 Spinacia Ol
C 51	19	2.0	2384	9	AK056891	AK056891 Homo sapi
C 52	19	2.0	2411	9	AK095799	AK095799 Homo sapi
C 53	19	2.0	2417	9	AK090755	AK090755 Homo sapi
C 54	19	2.0	2459	9	AK094806	AK094806 Homo sapi
C 55	19	2.0	2656	9	AK055499	AK055499 Homo sapi
C 56	19	2.0	2690	9	HSU07231	U07231 Homo sapien
C 57	19	2.0	2726	9	AK097055	AK097055 Homo sapi
C 58	19	2.0	4068	6	AX212239	AX212239 Sequence
C 59	19	2.0	4158	10	AB024336	AB024336 Mus muscu
C 60	19	2.0	21565	9	AL607068	AL607068 Human DNA
C 61	19	2.0	31312	9	AP000283	AP000283 Homo sapi
C 62	19	2.0	33140	6	AX067448	AX067448 Sequence
C 63	19	2.0	34968	3	U41534	U41534 Caenorhabdi
C 64	19	2.0	47403	6	AX059535	AX059535 Sequence
C 65	19	2.0	65970	9	AP000282	AP000282 Homo sapi

C 66	19	2.0	69513	2	AC131071	AC131071 Mus muscu
C 67	19	2.0	71272	2	AC103492	AC103492 Rattus no
C 68	19	2.0	73888	2	AC111426	AC111426 Rattus no
C 69	19	2.0	87286	2	AC014436	AC014436 Drosophil
C 70	19	2.0	91470	8	T4B21	AF118223 Arabidops
C 71	19	2.0	100000	9	AP000040	AP000040 Homo sapi
C 72	19	2.0	100000	9	AP000108	AP000108 Homo sapi
C 73	19	2.0	100000	9	AP000184	AP000184 Homo sapi
C 74	19	2.0	102196	9	AL590998	AL590998 Human DNA
C 75	19	2.0	109745	2	AC026447	AC026447 Homo sapi
C 76	19	2.0	110000	2	AC084053_2	Continuation (3 of
C 77	19	2.0	110000	2	AL359456_4	Continuation (5 of
C 78	19	2.0	110110	9	AC121337	AC121337 Homo sapi
C 79	19	2.0	111135	9	AL135933	AL135933 Human DNA
C 80	19	2.0	115915	9	AC105922	AC105922 Homo sapi
C 81	19	2.0	117743	2	AC109028	AC109028 Rattus no
C 82	19	2.0	119874	9	AL590096	AL590096 Human DNA
C 83	19	2.0	122467	9	AC129029	AC129029 Homo sapi
C 84	19	2.0	125108	9	AF030453	AF030453 Homo sapi
C 85	19	2.0	128532	2	AC128104	AC128104 Rattus no
C 86	19	2.0	135940	8	AC116603	AC116603 Oryza sat
C 87	19	2.0	137726	9	AC105919	AC105919 Homo sapi
C 88	19	2.0	140177	2	AC116245	AC116245 Rattus no
C 89	19	2.0	142504	2	AC129671	AC129671 Rattus no
C 90	19	2.0	143133	2	AC021655	AC021655 Homo sapi
C 91	19	2.0	148372	9	AC068388	AC068388 Homo sapi
C 92	19	2.0	150666	2	AC118817	AC118817 Rattus no
C 93	19	2.0	152185	2	AC012219	AC012219 Homo sapi
C 94	19	2.0	153619	2	AC130430	AC130430 Homo sapi
C 95	19	2.0	155182	2	AC023890	AC023890 Homo sapi
C 96	19	2.0	155221	9	AC016587	AC016587 Homo sapi
C 97	19	2.0	156264	2	AC102977	AC102977 Rattus no
C 98	19	2.0	159726	2	AC127949	AC127949 Rattus no
C 99	19	2.0	160380	9	AC112777	AC112777 Homo sapi
C 100	19	2.0	162354	9	AF260225	AF260225 Homo sapi
C 101	19	2.0	162876	3	AC008212	AC008212 Drosophil
C 102	19	2.0	164773	2	AC091543	AC091543 Felis cat
C 103	19	2.0	165933	8	AC131375	AC131375 Oryza sat
C 104	19	2.0	167439	2	AC130273	AC130273 Papio cyn
C 105	19	2.0	170914	3	AC010843	AC010843 Drosophil
C 106	19	2.0	171618	2	AC094532	AC094532 Rattus no
C 107	19	2.0	173067	9	AC083811	AC083811 Homo sapi
C 108	19	2.0	174033	2	AC021957	AC021957 Homo sapi
C 109	19	2.0	174084	2	AC084395	AC084395 Homo sapi
C 110	19	2.0	174311	9	AP002788	AP002788 Homo sapi
C 111	19	2.0	175034	2	AC130063	AC130063 Rattus no
C 112	19	2.0	176410	2	AC112712	AC112712 Homo sapi
C 113	19	2.0	176562	2	AC130272	AC130272 Papio cyn
C 114	19	2.0	179077	2	AC097075	AC097075 Rattus no
C 115	19	2.0	179323	2	AC011852	AC011852 Homo sapi
C 116	19	2.0	179591	9	AC093902	AC093902 Homo sapi
C 117	19	2.0	179693	2	AL691511	AL691511 Mus muscu
C 118	19	2.0	180069	9	AC007655	AC007655 Homo sapi
C 119	19	2.0	181191	9	AC117379	AC117379 Homo sapi
C 120	19	2.0	181700	2	AC128595	AC128595 Rattus no
C 121	19	2.0	181724	9	AC091983	AC091983 Homo sapi
C 122	19	2.0	181905	3	AC008211	AC008211 Drosophil
C 123	19	2.0	182087	2	AC055790	AC055790 Homo sapi
C 124	19	2.0	182087	2	AC055790	AC055790 Homo sapi
C 125	19	2.0	183784	2	AC101948	AC101948 Mus muscu
C 126	19	2.0	184001	9	AC104071	AC104071 Homo sapi
C 127	19	2.0	184692	2	AC094035	AC094035 Rattus no
C 128	19	2.0	185680	2	AC079738	AC079738 Homo sapi
C 129	19	2.0	185735	9	AC078785	AC078785 Homo sapi
C 130	19	2.0	185810	3	AC010995	AC010995 Drosophil
C 131	19	2.0	189534	2	AL354656	AL354656 Homo sapi
C 132	19	2.0	189999	9	AC092756	AC092756 Homo sapi
C 133	19	2.0	191137	2	AC124681	AC124681 Mus muscu
C 134	19	2.0	191275	2	AC015471	AC015471 Homo sapi
C 135	19	2.0	191440	2	AC095344	AC095344 Rattus no
C 136	19	2.0	192913	2	AC129214	AC129214 Mus muscu
C 137	19	2.0	193006	2	AC120599	AC120599 Rattus no
C 138	19	2.0	194872	10	AL590991	AL590991 Mouse DNA
C 139	19	2.0	196054	2	AC073388	AC073388 Mus muscu
C 140	19	2.0	198448	10	AC115118	AC115118 Mus muscu
C 141	19	2.0	200001	8	ATCHRIV13	AL161501 Arabidops
C 142	19	2.0	201805	10	AL606466	AL606466 Mouse DNA
C 143	19	2.0	201882	2	AC103345	AC103345 Rattus no
C 144	19	2.0	202972	10	AL663052	AL663052 Mouse DNA
C 145	19	2.0	203403	9	AC073130	AC073130 Homo sapi
C 146	19	2.0	205903	9	AP000901	AP000901 Homo sapi
C 147	19	2.0	207398	10	AL592462	AL592462 Mouse DNA
C 148	19	2.0	211902	3	AE003755	AE003755 Drosophil
C 149	19	2.0	215042	2	AC121594	AC121594 Mus muscu
C 150	19	2.0	219635	2	AL390207	AL390207 Homo sapi
C 151	19	2.0	238967	2	AC095377	AC095377 Rattus no
C 152	19	2.0	288763	2	AC124692	AC124692 Mus muscu
C 153	19	2.0	298839	3	AE003492	AE003492 Drosophil
C 154	19	2.0	305153	1	AP001520	AP001520 Bacillus
C 155	19	2.0	313573	3	CEY57G11C	Z99281 Caenorhabdi
C 156	19	2.0	316704	9	AF239258S3	AF223391 Homo sapi
C 157	19	2.0	337101	9	HSXKSRPXR	AL121578 Homo sapi
C 158	19	2.0	340000	9	AP001715	AP001715 Homo sapi
C 159	19	2.0	342800	1	AP003598	AP003598 Nostoc sp
C 160	18	1.9	803	9	AY077697	AY077697 Homo sapi
C 161	18	1.9	828	6	AX438187	AX438187 Sequence
C 162	18	1.9	1110	4	SSPGDMR	X16638 Pig pgd mRN
C 163	18	1.9	1268	10	CGU81045	U81045 Cricetulus
C 164	18	1.9	1362	3	EDHXX2	Y11115 E.dispar mR
C 165	18	1.9	1371	3	EDHXX1	Y11114 E.dispar mR
C 166	18	1.9	1372	3	EHXXK1	X82197 E.histolylti
C 167	18	1.9	1376	3	EHXXK2	X82198 E.histolylti
C 168	18	1.9	1440	3	AF525293	AF525293 Plasmodiu
C 169	18	1.9	1728	6	ARI39606	ARI39606 Sequence
C 170	18	1.9	1807	3	PFU07365	U07365 Plasmodium
C 171	18	1.9	1986	10	AF411816	AF411816 Mus muscu
C 172	18	1.9	3006	10	BC024690	BC024690 Mus muscu
C 173	18	1.9	4735	10	MUSMUC01	M76179 Mus musculu
C 174	18	1.9	5197	4	RABLS1B	D10108 Oryctolagus
C 175	18	1.9	6214	2	AC014611	AC014611 Drosophil
C 176	18	1.9	6833	5	XLWCYC	X53717 X.laervis c-
C 177	18	1.9	9051	2	AC014432	AC014432 Drosophil
C 178	18	1.9	10029	1	AE006274	AE006274 Lactococc
C 179	18	1.9	10078	1	AE006316	AE006316 Lactococc
C 180	18	1.9	10088	1	U32838	U32838 Haemophilus
C 181	18	1.9	10200	10	MMU16175	U16175 Mus musculu
C 182	18	1.9	10244	1	AE006251	AE006251 Lactococc
C 183	18	1.9	10330	1	AE006437	AE006437 Lactococc
C 184	18	1.9	10460	1	AE004248	AE004248 Vibrio ch
C 185	18	1.9	10775	1	AE006296	AE006296 Lactococc
C 186	18	1.9	11255	1	AE006443	AE006443 Lactococc
C 187	18	1.9	11393	1	AE006243	AE006243 Lactococc
C 188	18	1.9	14616	1	BSU09819	U09819 Bacillus su
C 189	18	1.9	18401	8	AY102584	AY102584 Gibberell
C 190	18	1.9	23688	3	U50310	U50310 Caenorhabdi
C 191	18	1.9	29371	1	AF164956	AF164956 Corynebac
C 192	18	1.9	29871	1	LLLPX214	X92946 Lactococcus
C 193	18	1.9	30113	9	HSCN19H5	Z72520 Human DNA s
C 194	18	1.9	30867	2	AC017487	AC017487 Drosophil
C 195	18	1.9	35554	9	HSDJ777D9	AL110120 Human DNA
C 196	18	1.9	35618	1	AF080002	AF080002 Hellobaci
C 197	18	1.9	35811	8	SC9718	Z49702 S.cerevisia
C 198	18	1.9	36948	2	AC014499	AC014499 Drosophil
C 199	18	1.9	40676	9	AC000088	AC000088 Homo sapi
C 200	18	1.9	43698	9	AC000071	AC000071 Homo sapi
C 201	18	1.9	50415	2	AC097426	AC097426 Rattus no
C 202	18	1.9	51960	8	AP004983	AP004983 Lotus jap
C 203	18	1.9	51984	2	AC182779	AC182779 Drosophil
C 204	18	1.9	52810	2	AC120420	AC120420 Mus muscu
C 205	18	1.9	53805	9	AC109360	AC109360 Homo sapi
C 206	18	1.9	57355	9	AL162720	AL162720 Human DNA
C 207	18	1.9	59903	2	AC015782	AC015782 Homo sapi
C 208	18	1.9	60481	2	AC101780	AC101780 Mus muscu
C 209	18	1.9	62047	2	AC115777	AC115777 Mus muscu
C 210	18	1.9	62934	2	AC112549	AC112549 Rattus no
C 211	18	1.9	63176	9	AL731571	AL731571 Human DNA

C 212	18	1.9	67145	8	AP004920	AP004920 Lotus jap	285	18	1.9	151806	8	AP002093	AP002093 Oryza sat
C 213	18	1.9	69070	2	AC100082	Mus muscu	C 286	18	1.9	152010	9	AC026785	AC026785 Homo sapi
C 214	18	1.9	69172	2	AC129960	Bos tauru	C 287	18	1.9	152238	2	AC119027	AC119027 Rattus no
C 215	18	1.9	72832	8	AC074228	Arabidops	C 288	18	1.9	152598	2	AC128741	AC128741 Rattus no
C 216	18	1.9	73021	2	AC017368	Arabidops	C 289	18	1.9	152702	9	AP000237	AP000237 Homo sapi
C 217	18	1.9	73391	8	AP002057	Arabidops	C 290	18	1.9	152896	2	AC110665	AC110665 Canis fam
C 218	18	1.9	78917	2	AC018146	Drosophil	C 291	18	1.9	152975	2	AC129220	AC129220 Mus muscu
C 219	18	1.9	81891	9	AC108076	Homo sapi	C 292	18	1.9	153160	2	AC110112	AC110112 Rattus no
C 220	18	1.9	83148	2	AC106381	Rattus no	C 293	18	1.9	153289	2	AC013533	AC013533 Homo sapi
C 221	18	1.9	85862	3	AC004433	Drosophil	C 294	18	1.9	153556	9	AC004926	AC004926 Homo sapi
C 222	18	1.9	87124	3	AC004371	Drosophil	C 295	18	1.9	153865	8	AP002744	AP002744 Oryza sat
C 223	18	1.9	89642	2	AC005136	AC005136 Homo sapi	C 296	18	1.9	153929	2	AC026023	AC026023 Homo sapi
C 224	18	1.9	90105	9	AL513007	Human DNA	C 297	18	1.9	153960	2	AC095558	AC095558 Rattus no
C 225	18	1.9	90170	2	AC098355	Rattus no	C 298	18	1.9	154311	9	AF107258	AF107258 Homo sapi
C 226	18	1.9	90347	5	AL714030	Zebrafish	C 299	18	1.9	154356	2	AC118518	AC118518 Rattus no
C 227	18	1.9	96660	9	AL442123	Human DNA	C 300	18	1.9	154597	9	AP004295	AP004295 Homo sapi
C 228	18	1.9	97348	10	AF091216	AF091216 Mus muscu	C 301	18	1.9	154793	2	AC016470	AC016470 Homo sapi
C 229	18	1.9	97433	2	AL512789	AL512789 Homo sapi	C 302	18	1.9	155424	2	AC129234	AC129234 Rattus no
C 230	18	1.9	98478	2	AC091370	AC091370 Rattus no	C 303	18	1.9	155852	2	AC115656	AC115656 Rattus no
C 231	18	1.9	100000	9	AP000093	AP000093 Homo sapi	C 304	18	1.9	156777	2	AC114972	AC114972 Homo sapi
C 232	18	1.9	100000	9	AP000197	AP000197 Homo sapi	C 305	18	1.9	158347	3	AC092217	AC092217 Drosophil
C 233	18	1.9	101097	2	AC127438	Rattus no	C 306	18	1.9	158434	9	AL500527	AL500527 Human DNA
C 234	18	1.9	102382	9	AC007557	AC007557 Homo sapi	C 307	18	1.9	158706	2	AC128748	AC128748 Rattus no
C 235	18	1.9	104125	3	AC084468	AC084468 Caenorhab	C 308	18	1.9	159096	9	AL161781	AL161781 Human DNA
C 236	18	1.9	106993	3	AE002751	AE002751 Drosophil	C 309	18	1.9	159693	10	AL646042	AL646042 Mouse DNA
C 237	18	1.9	108056	8	AC005170	AC005170 Arabidops	C 310	18	1.9	160460	2	AC024374	AC024374 Homo sapi
C 238	18	1.9	110000	2	AC074222	Mus muscu	C 311	18	1.9	161221	2	AC126711	AC126711 Rattus no
C 239	18	1.9	110000	2	AC111011	Continuation (2 of	C 312	18	1.9	161428	2	AC126925	AC126925 Canis fam
C 240	18	1.9	110000	2	AC112266	Continuation (2 of	C 313	18	1.9	161536	2	AC024318	AC024318 Homo sapi
C 241	18	1.9	110000	2	AC129176	Continuation (2 of	C 314	18	1.9	161834	2	AC114697	AC114697 Rattus no
C 242	18	1.9	110000	2	AL139235	Continuation (2 of	C 315	18	1.9	161873	8	AC090056	AC090056 Oryza sat
C 243	18	1.9	110000	2	AL139235	Continuation (3 of	C 316	18	1.9	162336	2	AC107170	AC107170 Rattus no
C 244	18	1.9	110000	2	AL596328	Continuation (2 of	C 317	18	1.9	162481	2	AC018954	AC018954 Homo sapi
C 245	18	1.9	110000	2	AL596328	Continuation (2 of	C 318	18	1.9	162680	9	AL158203	AL158203 Human DNA
C 246	18	1.9	111094	2	AP000683	AP000683 Homo sapi	C 319	18	1.9	162760	2	AC079229	AC079229 Homo sapi
C 247	18	1.9	112283	10	AL732574	AL732574 Mouse DNA	C 320	18	1.9	163012	3	AC008289	AC008289 Drosophil
C 248	18	1.9	113515	9	HSJ944FI3	AL117353 Human DNA	C 321	18	1.9	163641	2	AP005619	AP005619 Oryza sat
C 249	18	1.9	114741	2	AC111484	AC111484 Rattus no	C 322	18	1.9	163929	2	AC027686	AC027686 Homo sapi
C 250	18	1.9	116408	2	AC129743	AC129743 Rattus no	C 323	18	1.9	163958	3	AC010705	AC010705 Drosophil
C 251	18	1.9	120085	2	AC130016	AC130016 Rattus no	C 324	18	1.9	164035	3	AC007888	AC007888 Drosophil
C 252	18	1.9	120478	2	AC129844	AC129844 Rattus no	C 325	18	1.9	164049	2	AC126763	AC126763 Homo sapi
C 253	18	1.9	122947	2	AC106289	AC106289 Rattus no	C 326	18	1.9	164414	3	AC092395	AC092395 Drosophil
C 254	18	1.9	124048	9	HUAC002045	AC002045 Human Chr	C 327	18	1.9	164805	8	OSJN00174	AL662970 Oryza sat
C 255	18	1.9	125577	2	AC109541	AC109541 Rattus no	C 328	18	1.9	165693	9	AC087879	AC087879 Homo sapi
C 256	18	1.9	126227	2	AC103164	AC103164 Rattus no	C 329	18	1.9	165767	9	AC116917	AC116917 Homo sapi
C 257	18	1.9	129273	2	AC105320	AC105320 Oryza sat	C 330	18	1.9	165854	2	AC091539	AC091539 Canis fam
C 258	18	1.9	130005	9	AL590240	AL590240 Human DNA	C 331	18	1.9	166051	2	AC115126	AC115126 Rattus no
C 259	18	1.9	130755	2	AC093453	AC093453 Canis fam	C 332	18	1.9	166265	2	AC007841	AC007841 Homo sapi
C 260	18	1.9	131067	2	AC118536	AC118536 Canis fam	C 333	18	1.9	166368	2	AC127649	AC127649 Rattus no
C 261	18	1.9	132720	9	AC093269	AC093269 Homo sapi	C 334	18	1.9	166849	2	AC127842	AC127842 Rattus no
C 262	18	1.9	138261	2	AC1111740	AC111740 Rattus no	C 335	18	1.9	166869	9	AC016395	AC016395 Homo sapi
C 263	18	1.9	138490	10	MMU249895	AJ249895 Mus muscu	C 336	18	1.9	167480	9	HSBA18114	AL121928 Human DNA
C 264	18	1.9	138586	2	AC118891	AC118891 Rattus no	C 337	18	1.9	167499	2	CNS01DWX	AL138975 Homo sapi
C 265	18	1.9	139565	9	AL353786	AL353786 Human DNA	C 338	18	1.9	167981	2	AL591856	AL591856 Homo sapi
C 266	18	1.9	139649	9	AL359208	AL359208 Human DNA	C 339	18	1.9	168006	9	AC093740	AC093740 Homo sapi
C 267	18	1.9	139999	8	AC018727	AC018727 Oryza sat	C 340	18	1.9	168084	9	AC007037	AC007037 Homo sapi
C 268	18	1.9	140630	9	AL357514	AL357514 Human DNA	C 341	18	1.9	168438	2	AC024632	AC024632 Homo sapi
C 269	18	1.9	141756	9	AC024993	AC024993 Homo sapi	C 342	18	1.9	168602	2	AF271406	AF271406 Homo sapi
C 270	18	1.9	142123	9	AC004690	AC004690 Homo sapi	C 343	18	1.9	169146	2	AC101797	AC101797 Mus muscu
C 271	18	1.9	142408	2	AC126338	AC126338 Homo sapi	C 344	18	1.9	169340	9	AP001148	AP001148 Homo sapi
C 272	18	1.9	142616	2	AC067789	AC067789 Homo sapi	C 345	18	1.9	169460	4	AL731652	AL731652 Dog DNA s
C 273	18	1.9	144079	2	AC024342	AC024342 Homo sapi	C 346	18	1.9	169515	2	AC023213	AC023213 Homo sapi
C 274	18	1.9	145325	9	AL137789	AL137789 Human DNA	C 347	18	1.9	169649	2	AC046163	AC046163 Homo sapi
C 275	18	1.9	146582	10	AL662920	AL662920 Mouse DNA	C 348	18	1.9	170069	9	AC106814	AC106814 Homo sapi
C 276	18	1.9	147061	2	AC119597	AC119597 Rattus no	C 349	18	1.9	170138	2	AC018794	AC018794 Homo sapi
C 277	18	1.9	148246	8	AP003054	AP003054 Oryza sat	C 350	18	1.9	170939	2	AC128947	AC128947 Rattus no
C 278	18	1.9	148506	2	AC106345	AC106345 Rattus no	C 351	18	1.9	171029	2	AC096391	AC096391 Rattus no
C 279	18	1.9	148859	8	OSJN00093	AL606648 Oryza sat	C 352	18	1.9	171368	9	HSU95738	U95738 Human chrom
C 280	18	1.9	149078	2	AC117264	AC117264 Oryza sat	C 353	18	1.9	171681	2	AC091541	AC091541 Canis fam
C 281	18	1.9	149200	9	AL359645	AL359645 Human DNA	C 354	18	1.9	171700	2	AC094677	AC094677 Rattus no
C 282	18	1.9	149542	9	AC011711	AC011711 Homo sapi	C 355	18	1.9	171747	9	AP001554	AP001554 Homo sapi
C 283	18	1.9	149951	9	AC018375	AC018375 Homo sapi	C 356	18	1.9	171972	3	AC012388	AC012388 Drosophil
C 284	18	1.9	151721	9	AC013265	AC013265 Homo sapi	C 357	18	1.9	172367	2	AC118989	AC118989 Canis fam



C 358	18	1.9	173356	2	AP004047	AP004047 Oryza sat	431	18	1.9	199208	2	AC019236	AC019236 Homo sapi
C 359	18	1.9	173619	2	AC105962	AC105962 Mus muscu	C 432	18	1.9	200964	2	AL732553	AL732553 Mus muscu
C 360	18	1.9	173658	2	AC130996	AC130996 Rattus no	433	18	1.9	201410	9	HSP12	AL133500 Homo sapi
C 361	18	1.9	173898	9	AC026886	AC026886 Homo sapi	434	18	1.9	201437	2	AC015981	AC015981 Homo sapi
C 362	18	1.9	173904	9	AC025944	AC025944 Homo sapi	435	18	1.9	201803	2	AC113572	AC113572 Canis fam
C 363	18	1.9	174375	2	AC026220	AC026220 Homo sapi	C 436	18	1.9	202467	2	AL672232	AL672232 Mus muscu
C 364	18	1.9	174408	2	AC118165	AC118165 Rattus no	C 437	18	1.9	203281	2	AC126237	AC126237 Canis fam
C 365	18	1.9	175054	2	AC018915	AC018915 Homo sapi	C 438	18	1.9	203982	9	AC067854	AC067854 Homo sapi
C 366	18	1.9	175179	2	AC095321	AC095321 Rattus no	C 439	18	1.9	204062	2	AC091072	AC091072 Mus muscu
C 367	18	1.9	175426	2	AC113235	AC113235 Canis fam	C 440	18	1.9	204589	2	AC103655	AC103655 Mus muscu
C 368	18	1.9	176349	8	AP002092	AP002092 Oryza sat	441	18	1.9	204731	2	AC009714	AC009714 Homo sapi
C 369	18	1.9	176503	9	AL512604	AL512604 Human DNA	442	18	1.9	206160	2	AC113011	AC113011 Mus muscu
C 370	18	1.9	176840	2	AC115024	AC115024 Mus muscu	443	18	1.9	206463	2	AP000834	AP000834 Homo sapi
C 371	18	1.9	176840	2	AC105254	AC105254 Homo sapi	C 444	18	1.9	207336	2	AC103907	AC103907 Canis fam
C 372	18	1.9	177018	2	AC106357	AC106357 Rattus no	C 445	18	1.9	207698	9	AC114491	AC114491 Homo sapi
C 373	18	1.9	177108	2	AC026504	AC026504 Homo sapi	C 446	18	1.9	208548	2	AC092195	AC092195 Canis fam
C 374	18	1.9	177108	2	AC026504	AC026504 Homo sapi	447	18	1.9	208842	2	AC091295	AC091295 Mus muscu
C 375	18	1.9	177417	2	AC097924	AC097924 Rattus no	C 448	18	1.9	209904	10	AL611984	AL611984 Mouse DNA
C 376	18	1.9	177682	2	AC130050	AC130050 Rattus no	C 449	18	1.9	210496	10	AC104327	AC104327 Mus muscu
C 377	18	1.9	177918	10	AL593855	AL593855 Mouse DNA	C 450	18	1.9	210595	2	AC097034	AC097034 Rattus no
C 378	18	1.9	178067	2	AC130182	AC130182 Rattus no	C 451	18	1.9	211133	2	AC124522	AC124522 Mus muscu
C 379	18	1.9	179052	2	AC120232	AC120232 Rattus no	C 452	18	1.9	211860	2	AC113085	AC113085 Mus muscu
C 380	18	1.9	179596	2	AC094587	AC094587 Rattus no	453	18	1.9	212255	2	AC096705	AC096705 Rattus no
C 381	18	1.9	180257	2	AC034164	AC034164 Homo sapi	C 454	18	1.9	213030	2	AC094817	AC094817 Rattus no
C 382	18	1.9	180689	2	AC124913	AC124913 Sus scrof	C 455	18	1.9	213348	2	AC127212	AC127212 Rattus no
C 383	18	1.9	180827	9	AC007344	AC007344 Homo sapi	456	18	1.9	213401	10	AL671911	AL671911 Mouse DNA
C 384	18	1.9	180905	2	AC011688	AC011688 Homo sapi	C 457	18	1.9	214425	2	AC092091	AC092091 Canis fam
C 385	18	1.9	181021	2	AL845172	AL845172 Mus muscu	458	18	1.9	215286	2	AC111085	AC111085 Mus muscu
C 386	18	1.9	181065	2	AC114891	AC114891 Canis fam	459	18	1.9	216692	10	AL671906	AL671906 Mouse DNA
C 387	18	1.9	181542	2	AC121076	AC121076 Canis fam	C 460	18	1.9	218565	3	AC009257	AC009257 Drosophil
C 388	18	1.9	181930	2	AC125614	AC125614 Rattus no	C 461	18	1.9	218904	2	AL844840	AL844840 Mus muscu
C 389	18	1.9	182496	2	AC011190	AC011190 Homo sapi	C 462	18	1.9	220699	2	AC009206	AC009206 Drosophil
C 390	18	1.9	182594	2	AC121853	AC121853 Mus muscu	463	18	1.9	222061	2	AC105851	AC105851 Rattus no
C 391	18	1.9	182717	2	AC108399	AC108399 Mus muscu	464	18	1.9	222520	10	AC063968	AC063968 Genomic s
C 392	18	1.9	183722	3	AC009341	AC009341 Drosophil	465	18	1.9	223373	2	AC122881	AC122881 Mus muscu
C 393	18	1.9	184181	9	AL158141	AL158141 Human DNA	466	18	1.9	223624	2	AC093626	AC093626 Homo sapi
C 394	18	1.9	184722	2	AC110668	AC110668 Canis fam	467	18	1.9	224076	2	AC084019	AC084019 Mus muscu
C 395	18	1.9	184926	9	AP000757	AP000757 Homo sapi	C 468	18	1.9	225885	2	AC113236	AC113236 Canis fam
C 396	18	1.9	185587	2	AC103935	AC103935 Mus muscu	469	18	1.9	227979	9	AC008869	AC008869 Homo sapi
C 397	18	1.9	185666	2	AC048389	AC048389 Homo sapi	470	18	1.9	229961	2	AC122919	AC122919 Mus muscu
C 398	18	1.9	185930	2	AC109390	AC109390 Rattus no	C 471	18	1.9	230395	10	AL671878	AL671878 Mouse DNA
C 399	18	1.9	186844	2	AC090052	AC090052 Homo sapi	C 472	18	1.9	238169	2	AC091396	AC091396 Mus muscu
C 400	18	1.9	187349	9	AP003352	AP003352 Homo sapi	C 473	18	1.9	240294	2	AC044892	AC044892 Homo sapi
C 401	18	1.9	187908	9	HS495010	AL031121 Human DNA	C 474	18	1.9	246646	2	AL391275	AL391275 Homo sapi
C 402	18	1.9	188216	2	AC092711	AC092711 Mus muscu	C 475	18	1.9	251182	2	AC108230	AC108230 Rattus no
C 403	18	1.9	188674	2	AC114604	AC114604 Mus muscu	476	18	1.9	253950	2	AC125085	AC125085 Mus muscu
C 404	18	1.9	188854	10	AL626770	AL626770 Mouse DNA	477	18	1.9	254050	1	AL627269	AL627269 Salmonell
C 405	18	1.9	188992	2	AC097896	AC097896 Rattus no	478	18	1.9	259744	2	AL844181	AL844181 Mus muscu
C 406	18	1.9	189286	9	AC093906	AC093906 Homo sapi	C 479	18	1.9	261000	3	AE003654	AE003654 Drosophil
C 407	18	1.9	190082	2	AC120136	AC120136 Mus muscu	480	18	1.9	264191	2	AC106481	AC106481 Rattus no
C 408	18	1.9	190529	2	AC107732	AC107732 Mus muscu	481	18	1.9	264211	2	AC123867	AC123867 Mus muscu
C 409	18	1.9	190724	2	AC078877	AC078877 Homo sapi	C 482	18	1.9	267662	3	AE003577	AE003577 Drosophil
C 410	18	1.9	190743	2	AC073159	AC073159 Homo sapi	C 483	18	1.9	270852	2	AC125346	AC125346 Mus muscu
C 411	18	1.9	192044	9	AL590439	AL590439 Human DNA	484	18	1.9	275242	3	AE003649	AE003649 Drosophil
C 412	18	1.9	192235	2	AC025974	AC025974 Homo sapi	485	18	1.9	276746	2	AC126670	AC126670 Mus muscu
C 413	18	1.9	193044	10	AC117198	AC117198 Mus muscu	C 486	18	1.9	286963	2	AL732318	AL732318 Mus muscu
C 414	18	1.9	1933594	2	AC113281	AC113281 Mus muscu	C 487	18	1.9	295225	3	AE003461	AE003461 Drosophil
C 415	18	1.9	193638	2	AC101886	AC101886 Mus muscu	488	18	1.9	300600	1	AP005369	AP005369 Thermosyn
C 416	18	1.9	193641	10	AC121972	AC121972 Mus muscu	489	18	1.9	304247	3	AE003452	AE003452 Drosophil
C 417	18	1.9	193995	2	AC127254	AC127254 Mus muscu	490	18	1.9	306050	1	RME603645	AL603645 Rhizobium
C 418	18	1.9	194237	9	AC009063	AC009063 Homo sapi	491	18	1.9	306464	2	AC068708	AC068708 Homo sapi
C 419	18	1.9	194359	2	AC127720	AC127720 Rattus no	C 492	18	1.9	308092	3	AE003493	AE003493 Drosophil
C 420	18	1.9	194377	2	AC093120	AC093120 Mus muscu	493	18	1.9	337178	3	DROSADH08	AE003414 Drosophil
C 421	18	1.9	194602	9	AC113188	AC113188 Homo sapi	494	18	1.9	340000	9	AP001706	AP001706 Homo sapi
C 422	18	1.9	195224	2	AC120767	AC120767 Rattus no	C 495	17	1.7	111	9	HS39E8R	Z60932 H.sapiens C
C 423	18	1.9	195794	2	AC114924	AC114924 Mus muscu	C 496	17	1.7	117	6	A06329	A06329 Hpv18 gene
C 424	18	1.9	196555	2	AC092882	AC092882 Homo sapi	C 497	17	1.7	186	12	SYNAFPG	M59842 Synthetic f
C 425	18	1.9	196715	2	AC094874	AC094874 Rattus no	498	17	1.7	193	6	ARI04631	ARI04631 Sequence
C 426	18	1.9	197348	2	AC097095	AC097095 Rattus no	499	17	1.7	193	6	I38214	I38214 Sequence 44
C 427	18	1.9	197550	9	CNS01RHP	ALI62171 Human chr	C 500	17	1.7	344	6	AX208916	AX208916 Sequence
C 428	18	1.9	197870	2	AC091075	AC091075 Homo sapi	C 501	17	1.7	344	6	AX209654	AX209654 Sequence
C 429	18	1.9	198791	2	AC102542	AC102542 Mus muscu	C 502	17	1.7	382	11	G42550	G42550 SHGC-58812
C 430	18	1.9	199046	2	AC126106	AC126106 Rattus no	C 503	17	1.7	387	11	G47478	G47478 Z24974_1 Ze



504	17	1.7	402	6	AR137574	AR137574 Sequence	577	17	1.7	2178	5	AB008162	AB008162 Xenopus l
505	17	1.7	402	6	AR139080	AR139080 Sequence	578	17	1.7	2195	9	HSU50436	U50436 Human Down
506	17	1.7	421	11	AF303455	AF303455 Drosophila	C 579	17	1.7	2325	10	AF155196	AF155196 Rattus no
C 507	17	1.7	449	6	AX198361	AX198361 Sequence	C 580	17	1.7	2327	9	HSU50428	U50428 Human Down
508	17	1.7	477	5	ZEFMHDABAI	L04813 Brachydanio	581	17	1.7	2387	6	I25974	I25974 Sequence 4
C 509	17	1.7	479	11	AF288327	AF288327 Drosophila	582	17	1.7	2428	9	HSU50439	U50439 Human Down
510	17	1.7	612	6	AX066293	AX066293 Sequence	583	17	1.7	2469	10	AF013274	AF013274 Mus muscu
511	17	1.7	673	10	RNCEH	X60328 R.norvegicu	C 584	17	1.7	2566	9	HSU43378	U43378 Human Down
512	17	1.7	684	10	AF120322	AF120322 Mus muscu	585	17	1.7	2586	10	MMIFNAR2B	Y09864 M.musculus
C 513	17	1.7	734	9	HSA328094	AJ328094 Homo sapi	586	17	1.7	2659	10	AF440786	AF440786 Mus muscu
514	17	1.7	747	6	AX053321	AX053321 Sequence	587	17	1.7	2715	10	MMIFNAR2C	Y09865 M.musculus
515	17	1.7	747	6	AX053430	AX053430 Sequence	C 588	17	1.7	2742	4	BOV2PF36G	L42319 Bos taurus
C 516	17	1.7	756	8	SCYKR083C	Z28308 S.cerevisia	C 589	17	1.7	2884	8	TOVHMG2A	M63642 Tomato 3-hy
517	17	1.7	830	9	HSA334408	AJ334408 Homo sapi	590	17	1.7	2925	1	PACIOAB	Y10528 P.aeruginos
C 518	17	1.7	882	10	AF290952	AF290952 Mus muscu	C 591	17	1.7	2965	14	AB067537	AB067537 Norwalk v
C 519	17	1.7	942	10	AF290953	AF290953 Mus muscu	C 592	17	1.7	2965	14	AB067538	AB067538 Norwalk v
C 520	17	1.7	963	3	AF205376	AF205376 Plasmodiu	C 593	17	1.7	3019	1	AF270315	AF270315 Staphyloc
C 521	17	1.7	967	11	CNS06EIL	AL395155 T7 end of	C 594	17	1.7	3019	6	AX145633	AX145633 Sequence
522	17	1.7	969	6	AX142689	AX142689 Sequence	C 595	17	1.7	3078	3	AF318613	AF318613 Caenorhab
523	17	1.7	969	6	AX143661	AX143661 Sequence	C 596	17	1.7	3129	9	AF077599	AF077599 Homo sapi
C 524	17	1.7	987	10	AF290954	AF290954 Mus muscu	597	17	1.7	3532	9	AK027409	AK027409 Homo sapi
525	17	1.7	1002	6	E14040	E14040 CDNA encodi	C 598	17	1.7	3586	8	DGR288901	AJ288901 Dendrobii
C 526	17	1.7	1104	14	PNE306820	AJ306820 Prunus ne	C 599	17	1.7	3591	8	AB032182	AB032182 Physcomit
C 527	17	1.7	1105	14	PNE306822	AJ306822 Prunus ne	600	17	1.7	3720	5	XELNCAM	M25696 X.laavis ne
C 528	17	1.7	1106	14	PNE306812	AJ306812 Prunus ne	601	17	1.7	3720	5	AX339061	AX339061 Sequence
C 529	17	1.7	1106	14	PNE306824	AJ306824 Prunus ne	602	17	1.7	3799	6	AF092564	AF092564 Homo sapi
C 530	17	1.7	1112	6	I15087	I15087 Sequence 1	603	17	1.7	3799	9	AF092564	AF092564 Homo sapi
C 531	17	1.7	1113	3	PFARR2	L22058 Plasmodium	C 604	17	1.7	3831	2	AC017735	AC017735 Drosophil
532	17	1.7	1113	5	PGNLDHC4A	L79958 Columba liv	C 605	17	1.7	3994	9	AL772251	AL772251 Human DNA
533	17	1.7	1146	3	AF404110	AF404110 Junonia c	606	17	1.7	4082	6	AX440472	AX440472 Sequence
534	17	1.7	1185	5	DUKLDHBCRY	J03869 Duck lactat	607	17	1.7	4086	9	AB019987	AB019987 Homo sapi
535	17	1.7	1209	5	XELRDS38A	L79915 Xenopus lae	608	17	1.7	4093	1	AF269753	AF269753 Staphyloc
C 536	17	1.7	1233	10	AF290955	AF290955 Mus muscu	609	17	1.7	4093	6	AX145071	AX145071 Sequence
537	17	1.7	1234	5	AF069771	AF069771 Gallus ga	610	17	1.7	4217	8	SPRAD3	X63544 S.pombe rad
538	17	1.7	1240	6	AR050563	AR050563 Sequence	611	17	1.7	4267	8	SCYJLH092W	Z49367 S.cerevisia
539	17	1.7	1242	5	L79957	L79957 Columba liv	612	17	1.7	4660	5	APLDHBD	X68811 A.platythrnc
540	17	1.7	1254	8	AF002691	AF002691 Podocarpu	613	17	1.7	4660	8	SCRADH	X15665 Yeast RADH
541	17	1.7	1267	4	SSU95378	U95378 Sus scrofa	C 614	17	1.7	4881	6	BD003805	BD003805 Polynucle
542	17	1.7	1294	5	L79954	L79954 Trachemys s	615	17	1.7	5042	9	HSM805132	AL833949 Homo sapi
543	17	1.7	1295	5	AF363795	AF363795 Pelodiscu	616	17	1.7	5107	6	AX345968	AX345968 Sequence
544	17	1.7	1344	6	AX123416	AX123416 Sequence	C 617	17	1.7	5210	14	HPU893349	U89349 Human papil
C 545	17	1.7	1345	9	HSM800836	AL110186 Homo sapi	618	17	1.7	5245	1	CEYNNMOB	L49438 Chelatobact
C 546	17	1.7	1362	10	AF290956	AF290956 Mus muscu	619	17	1.7	5261	9	HSM801845	AL136877 Homo sapi
547	17	1.7	1370	5	AF218799	AF218799 Gallus ga	C 620	17	1.7	5327	9	AY028807	AY028807 Homo sapi
C 548	17	1.7	1373	3	PFU01322	U01322 Plasmodium	C 621	17	1.7	5416	10	MMU223996	AJ223956 Mus muscu
549	17	1.7	1389	5	XLNLPNO38	X56039 X.laavis mr	C 622	17	1.7	5600	6	AR124758	AR124758 Sequence
550	17	1.7	1404	6	AR120242	AR120242 Sequence	623	17	1.7	5857	1	NMPDHC	X82637 N.meningiti
C 551	17	1.7	1404	10	MUSINVO	L28819 Mus musculu	C 624	17	1.7	5932	1	CHU39411	U39411 Chelatobact
C 552	17	1.7	1421	6	AX452915	AX452915 Sequence	C 625	17	1.7	6056	2	AC017797	AC017797 Drosophil
C 553	17	1.7	1524	6	AR047985	AR047985 Sequence	C 626	17	1.7	6528	9	AB011097	AB011097 Homo sapi
C 554	17	1.7	1524	6	AR059325	AR059325 Sequence	C 627	17	1.7	6786	10	MMNGA28S	X86000 Mus musculu
C 555	17	1.7	1528	8	AF152200	AF152200 Libocedru	C 628	17	1.7	7022	10	MMPOLAS6	Y09488 M.musculus
C 556	17	1.7	1530	8	AF152177	AF152177 Austroced	C 629	17	1.7	7035	3	CELCYT1A	L26545 Caenorhabdi
557	17	1.7	1552	3	AY051510	AY051510 Drosophil	630	17	1.7	7296	5	AF133907	AF133907 Anolis ca
C 558	17	1.7	1620	10	AF290957	AF290957 Mus muscu	631	17	1.7	7491	6	AX346487	AX346487 Sequence
C 559	17	1.7	1638	8	AF351191	AF351191 Zea mays	C 632	17	1.7	7650	2	AC018280	AC018280 Drosophil
C 560	17	1.7	1640	9	AB070034	AB070034 Macaca fa	C 633	17	1.7	7829	1	RPPLAS	X71612 Buchnera ap
C 561	17	1.7	1646	14	AF034991	AF034991 Prunus ne	C 634	17	1.7	7857	14	PAPHPV18	X05015 Human papil
C 562	17	1.7	1652	14	AF034992	AF034992 Prunus ne	C 635	17	1.7	7940	9	HSM804873	AL833560 Homo sapi
C 563	17	1.7	1653	14	AF034993	AF034993 Prunus ne	C 636	17	1.7	8022	6	A61387	A61387 Sequence 3
C 564	17	1.7	1686	14	AF013287	AF013287 Prunus ne	C 637	17	1.7	8022	8	SPRAD3GEN	Y09076 S.pombe RAD
565	17	1.7	1756	9	AK002200	AK002200 Homo sapi	C 638	17	1.7	8022	8	SPU76307	U76307 Schizosacch
566	17	1.7	1786	3	PFTBPM	X77914 P.falciapar	C 639	17	1.7	8226	8	SPU66305	U66305 Schizosacch
C 567	17	1.7	1873	9	HSM804904	AL833591 Homo sapi	640	17	1.7	8615	1	AE006711	AE006711 Sulfolobu
568	17	1.7	1950	3	AY095004	AY095004 Drosophil	641	17	1.7	9354	10	RNO18435	Y18435 Rattus norv
C 569	17	1.7	1951	14	PNRVRNA3	Y07568 Prunus necr	C 642	17	1.7	9464	8	SJ100567	X77087 S.cerevisia
570	17	1.7	1954	6	AR171283	AR171283 Sequence	C 643	17	1.7	9782	1	AE009941	AE009941 Pyrobacul
C 571	17	1.7	1957	14	PNU57046	U57046 Prunus necr	C 644	17	1.7	10029	1	AE013368	AE013368 Methanosa
C 572	17	1.7	1991	9	BC003517	BC003517 Homo sapi	C 645	17	1.7	10299	1	AE004585	AE004585 Pseudomon
573	17	1.7	1992	6	AX402040	AX402040 Sequence	646	17	1.7	10311	1	AE010974	AE010974 Methanosa
574	17	1.7	1992	10	RNCEHR	X65083 R.norvegicu	647	17	1.7	10605	1	AE013050	AE013050 Thermoana
575	17	1.7	2023	10	MMU10484	U10484 Mus musculu	C 648	17	1.7	11342	1	AE003882	AE003882 Xylella f
576	17	1.7	2122	9	HSM800609	AL080105 Homo sapi	649	17	1.7	11353	1	AE007326	AE007326 Streptoco

C 650	17	1.7	11925	9	AC023596	AC023596 Homo sapi	17	1.7	60555	2	AC100824	AC100824 Homo sapi
651	17	1.7	11968	1	AE002482	AE002482 Neisseria	17	1.7	60676	2	AC084711	AC084711 Homo sapi
C 652	17	1.7	11976	1	AE006249	AE006249 Lactococc	17	1.7	60815	2	AC100876	AC100876 Mus muscu
C 653	17	1.7	12119	1	AE004810	AE004810 Pseudomon	17	1.7	61204	2	AC017132	AC017132 Drosophil
654	17	1.7	12280	1	AE006282	AE006282 Lactococc	17	1.7	62214	2	OSIG00034	AL732336 Oryza sat
655	17	1.7	13030	1	AE008392	AE008392 Streptoco	17	1.7	62421	8	NCB3E4	AL355931 Neurospor
C 656	17	1.7	13186	1	AE0001613	AE001613 Chlamydia	17	1.7	62953	2	AC100701	AC100701 Mus muscu
657	17	1.7	13527	1	AE010780	AE010780 Methanosa	17	1.7	63170	10	AL645799	AL645799 Mouse DNA
658	17	1.7	14494	1	AE010475	AE010475 Fusobacte	17	1.7	63618	2	AC129959	AC129959 Bos tauru
C 659	17	1.7	14960	1	AVU83986	U83986 Agrobacteri	17	1.7	64011	9	AC007642	AC007642 Homo sapi
660	17	1.7	15758	2	AC014859	AC014859 Drosophil	17	1.7	64064	2	AC117767	AC117767 Mus muscu
C 661	17	1.7	16310	10	RNO428514	AJ428514 Rattus no	17	1.7	64478	9	AL603839	AL603839 Human DNA
662	17	1.7	16535	2	AC014950	AC014950 Drosophil	17	1.7	64569	8	AB026650	AB026650 Arabidops
663	17	1.7	16652	2	AC018070	AC018070 Drosophil	17	1.7	64792	9	AL390028	AL390028 Human DNA
C 664	17	1.7	16720	8	LEU68072	U68072 Lycopersico	17	1.7	64855	2	AC015482	AC015482 Homo sapi
665	17	1.7	16766	6	AX347059	AX347059 Sequence	17	1.7	64873	9	AC108071	AC108071 Homo sapi
C 666	17	1.7	17527	1	AE000098	AE000098 Rhizobium	17	1.7	64942	2	AC117679	AC117679 Mus muscu
C 667	17	1.7	18002	8	SCUNORF1	Z27116 S.cerevisia	17	1.7	65034	2	AC016067	AC016067 Homo sapi
668	17	1.7	20369	9	AF026276	AF026276 Homo sapi	17	1.7	65395	2	AC126286	AC126286 Rattus no
669	17	1.7	21783	1	AE002208	AE002208 Chlamydog	17	1.7	66585	9	AL512378	AL512378 Human DNA
670	17	1.7	23306	1	AE008747	AE008747 Salmonell	17	1.7	66668	2	AC102913	AC102913 Mus muscu
671	17	1.7	23645	8	SPBC2G5	AL033385 S.pombe c	17	1.7	66857	2	AC068239	AC068239 Homo sapi
672	17	1.7	23919	3	AE003091	AE003091 Drosophil	17	1.7	67201	9	AL136169	AL136169 Human DNA
673	17	1.7	24206	9	AC128658	AC128658 Homo sapi	17	1.7	67434	2	AC101255	AC101255 Mus muscu
674	17	1.7	25311	8	SPBC216	AL049558 S.pombe c	17	1.7	67960	2	AC118695	AC118695 Mus muscu
675	17	1.7	25416	2	AC013100	AC013100 Drosophil	17	1.7	68507	2	AC126323	AC126323 Homo sapi
676	17	1.7	26411	3	U23486	U23486 Caenorhabdi	17	1.7	68974	2	AC074063	AC074063 Mus muscu
677	17	1.7	28573	3	AF024502	AF024502 Caenorhab	17	1.7	69096	2	AC121147	AC121147 Mus muscu
C 678	17	1.7	28984	9	AB002059	AB002059 Homo sapi	17	1.7	69226	2	AC118209	AC118209 Mus muscu
679	17	1.7	29034	9	HSA243947	AJ243947 Homo sapi	17	1.7	69373	2	AC101456	AC101456 Mus muscu
680	17	1.7	29158	9	AC119739	AC119739 Homo sapi	17	1.7	69458	9	AC025432	AC025432 Homo sapi
681	17	1.7	29364	3	CEC27B7	Z54236 Caenorhabdi	17	1.7	69646	2	AC090462	AC090462 Homo sapi
C 682	17	1.7	29498	8	AB023029	AB023029 Arabidops	17	1.7	69793	9	AC104646	AC104646 Homo sapi
C 683	17	1.7	29552	3	U64845	U64845 Caenorhabdi	17	1.7	69880	2	AC099820	AC099820 Homo sapi
684	17	1.7	31132	8	AB026640	AB026640 Arabidops	17	1.7	69986	2	AC020283	AC020283 Drosophil
C 685	17	1.7	31441	9	AL592163	AL592163 Human DNA	17	1.7	70820	2	AC017451	AC017451 Drosophil
C 686	17	1.7	31469	3	AF026211	AF026211 Caenorhab	17	1.7	71184	8	AP001298	AP001298 Arabidops
687	17	1.7	33477	3	CET07C4	Z29443 Caenorhabdi	17	1.7	72137	9	AL356795	AL356795 Human DNA
688	17	1.7	33866	2	AC125312	AC125312 Mus muscu	17	1.7	72198	2	AC123696	AC123696 Mus muscu
689	17	1.7	36120	3	U39997	U39997 Caenorhabdi	17	1.7	73131	2	AC016443	AC016443 Homo sapi
C 690	17	1.7	36949	7	AF323670	AF323670 Bacteriop	17	1.7	73509	2	AL137862	AL137862 Homo sapi
C 691	17	1.7	37482	3	AC024790	AC024790 Caenorhab	17	1.7	73656	2	AC016512	AC016512 Homo sapi
692	17	1.7	38393	9	HSN62C4	Z73963 Human DNA s	17	1.7	73999	8	AB019226	AB019226 Arabidops
C 693	17	1.7	38490	9	AP001236	AP001236 Homo sapi	17	1.7	74520	9	AP001264	AP001264 Homo sapi
694	17	1.7	38746	2	AC014782	AC014782 Drosophil	17	1.7	75163	2	AC069103	AC069103 Homo sapi
C 695	17	1.7	38969	3	U51995	U51995 Caenorhabdi	17	1.7	75245	2	AC129847	AC129847 Rattus no
C 696	17	1.7	39577	9	AL590671	AL590671 Human DNA	17	1.7	75647	2	AC101442	AC101442 Mus muscu
C 697	17	1.7	40885	2	AC100188	AC100188 Mus muscu	17	1.7	76299	2	AC019585	AC019585 Drosophil
698	17	1.7	41194	10	AF367979	AF367979 Mus muscu	17	1.7	77547	2	AC016804	AC016804 Homo sapi
699	17	1.7	41623	9	HSMQ16F8	Z72006 Human DNA s	17	1.7	77547	2	AC016804	AC016804 Homo sapi
700	17	1.7	42338	9	HUMQ16F8	D85922 Human DNA f	17	1.7	79048	2	AC024976	AC024976 Homo sapi
701	17	1.7	42854	9	AP001242	AP001242 Homo sapi	17	1.7	80204	2	AC016191	AC016191 Homo sapi
C 702	17	1.7	46208	2	AL353712	AL353712 Homo sapi	17	1.7	80514	9	AC004454	AC004454 Homo sapi
C 703	17	1.7	46864	2	AC100185	AC100185 Mus muscu	17	1.7	81195	10	AC020807	AC020807 Mus muscu
704	17	1.7	48508	9	AC084167	AC084167 Homo sapi	17	1.7	81195	10	AC091750	AC091750 Mus muscu
705	17	1.7	51305	9	AL691465	AL691465 Human DNA	17	1.7	81255	2	AC017607	AC017607 Drosophil
706	17	1.7	51556	9	AL160071	AL160071 Human DNA	17	1.7	81587	3	AF003135	AF003135 Caenorhab
707	17	1.7	51575	8	AP004492	AP004492 Lotus jap	17	1.7	81672	8	AB020755	AB020755 Arabidops
C 708	17	1.7	51757	9	AC007025	AC007025 Homo sapi	17	1.7	83823	2	AC006789	AC006789 Caenorhab
C 709	17	1.7	52707	2	AC017288	AC017288 Drosophil	17	1.7	84234	9	AC107975	AC107975 Homo sapi
710	17	1.7	53907	2	AC098773_4	Continuation (5 of	17	1.7	84392	2	AC013393	AC013393 Homo sapi
C 711	17	1.7	54527	2	AC1015174_3	AC015174 Drosophil	17	1.7	84997	2	AC123571	AC123571 Medicago
C 712	17	1.7	55252	2	AC119489_3	Continuation (4 of	17	1.7	85432	9	AC099799	AC099799 Homo sapi
713	17	1.7	55956	2	AC101559	AC101559 Mus muscu	17	1.7	85710	8	ATF28O9	AL137080 Arabidops
714	17	1.7	56733	2	AC017772	AC017772 Drosophil	17	1.7	85713	9	HSY237C10_3	Continuation (4 of
715	17	1.7	56954	9	AL590663	AL590663 Human DNA	17	1.7	86298	9	AL355347	AL355347 Human DNA
C 716	17	1.7	57623	2	AC105212	AC105212 Homo sapi	17	1.7	86332	9	AL445470	AL445470 Human DNA
717	17	1.7	57983	2	AC119877	AC119877 Mus muscu	17	1.7	86723	9	AF401203	AF401203 Homo sapi
718	17	1.7	58433	2	AC100304	AC100304 Mus muscu	17	1.7	87229	2	AC123538	AC123538 Sminthops
719	17	1.7	58641	2	AC104030	AC104030 Homo sapi	17	1.7	87229	2	AC123538	AC123538 Sminthops
C 720	17	1.7	59261	8	T12M4	AC003114 Arabidops	17	1.7	87286	2	AC014436	AC014436 Drosophil
C 721	17	1.7	59349	2	AC094968	AC094968 Rattus no	17	1.7	88413	8	AC005724	AC005724 Arabidops
C 722	17	1.7	59449	2	AC101666	AC101666 Mus muscu	17	1.7	89297	10	AL606920	AL606920 Mouse DNA

796	17	1.7	89354	9	AC008957	AC008957 Homo sapi	869	17	1.7	112662	9	AC104453	AC104453 Homo sapi
797	17	1.7	89862	9	AC007164	AC007164 Homo sapi	c 870	17	1.7	113284	30	AC074300	AC074300 Homo sapi
798	17	1.7	90289	2	AP003164	AP003164 Homo sapi	c 871	17	1.7	113350	2	AC120244	AC120244 Rattus no
799	17	1.7	91308	2	AC098294	AC098294 Rattus no	c 872	17	1.7	113546	2	HSAC000382	AC000382 Homo sapi
800	17	1.7	91708	2	AC123950_3	Continuation (4 of	873	17	1.7	113951	9	AC007552	AC007552 Homo sapi
c 801	17	1.7	92143	9	HS608E8	AL022343 Human DNA	874	17	1.7	114279	2	AC083907	AC083907 Homo sapi
802	17	1.7	92458	9	AL353715	AL353715 Human DNA	875	17	1.7	114414	9	AC073073	AC073073 Homo sapi
803	17	1.7	93432	9	CNS01DSR	AL121840 Human chr	876	17	1.7	114498	8	F3O9	AC006341 Arabidops
c 804	17	1.7	93485	8	AC011808	AC011808 Arabidops	877	17	1.7	114656	8	ATT15B3	AL163975 Arabidops
805	17	1.7	93645	8	AP004954	AP004954 Lotus jap	878	17	1.7	114793	9	AF217796	AF217796 Homo sapi
806	17	1.7	94128	2	AC131393	AC131393 Homo sapi	879	17	1.7	115000	9	AC104831	AC104831 Homo sapi
c 807	17	1.7	94319	9	HS341110	Z97352 Human DNA s	880	17	1.7	115468	2	AC011820	AC011820 Homo sapi
808	17	1.7	95298	9	AL365398	AL365398 Human DNA	c 881	17	1.7	115568	9	AC092120	AC092120 Homo sapi
c 809	17	1.7	96526	2	AC022385	AC022385 Homo sapi	c 882	17	1.7	115721	8	F25A4	AC008263 Arabidops
c 810	17	1.7	96640	2	AC106184	AC106184 Rattus no	883	17	1.7	115958	2	AC129038	AC129038 Rattus no
c 811	17	1.7	96671	2	AC097401	AC097401 Rattus no	884	17	1.7	117273	2	AC015138	AC015138 Drosophil
c 812	17	1.7	96764	9	AL732325	AL732325 Human DNA	885	17	1.7	118235	10	AC003060	AC003060 Mus muscu
813	17	1.7	96879	9	HS384F21	AL022171 Human DNA	c 886	17	1.7	118780	2	AC097699	AC097699 Rattus no
814	17	1.7	97057	9	AL627344	AL627344 Human DNA	887	17	1.7	119028	2	AC098832	AC098832 Oryza sat
c 815	17	1.7	97263	8	AC013258	AC013258 Arabidops	c 888	17	1.7	119243	2	AC112576	AC112576 Rattus no
c 816	17	1.7	97860	9	AC091816	AC091816 Homo sapi	c 889	17	1.7	119330	9	AL590076	AL590076 Human DNA
c 817	17	1.7	97912	9	HSJ753D10	AL049651 Human DNA	c 890	17	1.7	119562	2	AC006779	AC006779 Caenorhab
c 818	17	1.7	97959	9	AC093085	AC093085 Homo sapi	891	17	1.7	119821	2	CNS07YP6	AL713929 Oryza sat
819	17	1.7	98351	2	AC111997	AC111997 Rattus no	c 892	17	1.7	119972	2	AP004029	AP004029 Oryza sat
c 820	17	1.7	98569	9	HS209B9	AL021879 Human DNA	893	17	1.7	120000	2	HSAC000001	AC000001 Homo sapi
821	17	1.7	98915	2	AC108285	AC108285 Rattus no	c 894	17	1.7	120512	9	AL359852	AL359852 Human DNA
c 822	17	1.7	99072	2	AC130096	AC130096 Rattus no	895	17	1.7	120724	9	CNS05TCR	AL355837 Human chr
823	17	1.7	99084	2	AC026420	AC026420 Homo sapi	c 896	17	1.7	120780	2	AC095467	AC095467 Rattus no
824	17	1.7	99163	9	AL355602	AL355602 Human DNA	c 897	17	1.7	120834	9	AC093602	AC093602 Homo sapi
825	17	1.7	100307	2	AC111696	AC111696 Rattus no	898	17	1.7	120864	2	AC128843	AC128843 Rattus no
c 826	17	1.7	100785	2	AC069169	AC069169 Homo sapi	899	17	1.7	121027	2	AL157891	AL157891 Homo sapi
c 827	17	1.7	100843	2	AC094430	AC094430 Rattus no	c 900	17	1.7	121151	9	HS705D16	AL034428 Human DNA
c 828	17	1.7	100953	9	AL596269	AL596269 Homo sapi	901	17	1.7	121162	6	AX039602	AX039602 Sequence
c 829	17	1.7	101427	9	AL596269	AL596269 Homo sapi	902	17	1.7	122073	2	AC090831	AC090831 Homo sapi
c 830	17	1.7	101574	9	HS206D15	AL021068 Human DNA	c 903	17	1.7	122511	2	AC118494	AC118494 Rattus no
831	17	1.7	101635	2	AC097932	AC097932 Rattus no	904	17	1.7	122697	9	AC006062	AC006062 Homo sapi
832	17	1.7	102735	5	AL606702	AL606702 Zebrafish	c 905	17	1.7	122844	2	H0719D04	AL442106 Oryza sat
833	17	1.7	102794	9	AC104623	AC104623 Homo sapi	906	17	1.7	123369	9	AC025170	AC025170 Homo sapi
834	17	1.7	102822	9	AC104407	AC104407 Homo sapi	c 907	17	1.7	123436	9	AC009872	AC009872 Homo sapi
835	17	1.7	102848	2	AC128275	AC128275 Rattus no	c 908	17	1.7	123764	2	OSJN00097	AL606650 Oryza sat
836	17	1.7	102873	8	ATF26G5	AL353814 Arabidops	909	17	1.7	124040	2	AC097412	AC097412 Rattus no
c 837	17	1.7	103211	2	AC106975	AC106975 Rattus no	910	17	1.7	124114	9	HSJ858M22	AL118510 Human DNA
c 838	17	1.7	103223	8	AC007153	AC007153 Arabidops	c 911	17	1.7	124696	2	AP000649	AP000649 Homo sapi
c 839	17	1.7	103718	9	AL139175	AL139175 Human DNA	c 912	17	1.7	124912	2	AC129811	AC129811 Rattus no
840	17	1.7	103749	9	AL355810	AL355810 Human DNA	913	17	1.7	125086	9	AL607044	AL607044 Human DNA
c 841	17	1.7	104129	2	AC128003	AC128003 Rattus no	c 914	17	1.7	125198	2	AC098410	AC098410 Rattus no
c 842	17	1.7	104395	9	AC123021	AC123021 Homo sapi	c 915	17	1.7	125346	2	AP004329	AP004329 Oryza sat
c 843	17	1.7	104895	2	AC093288	AC093288 Homo sapi	916	17	1.7	126038	8	AP000367	AP000367 Oryza sat
c 844	17	1.7	105100	9	HS67A5	AL035594 Human DNA	917	17	1.7	126046	2	AC103657	AC103657 Mus muscu
c 845	17	1.7	106684	5	AL672107	AL672107 Zebrafish	c 918	17	1.7	126046	2	AC103657	AC103657 Mus muscu
c 846	17	1.7	106928	2	AC111925	AC111925 Rattus no	c 919	17	1.7	126707	2	AC123141	AC123141 Rattus no
c 847	17	1.7	108449	2	AC121711	AC121711 Rattus no	c 920	17	1.7	126779	9	AL603882	AL603882 Human DNA
848	17	1.7	108456	2	AC094316	AC094316 Rattus no	921	17	1.7	126780	2	AL138904	AL138904 Homo sapi
c 849	17	1.7	108841	9	AL359185	AL359185 Human DNA	922	17	1.7	127119	2	AC123277	AC123277 Rattus no
850	17	1.7	109288	9	AL359081	AL359081 Human DNA	923	17	1.7	127399	9	AL158825	AL158825 Human DNA
c 851	17	1.7	109757	9	AL162379	AL162379 Human DNA	924	17	1.7	127439	9	AC126178	AC126178 Homo sapi
c 852	17	1.7	110000	2	AC091454_1	Continuation (2 of	c 925	17	1.7	127713	2	OSJN00007	AL606452 Oryza sat
c 853	17	1.7	110000	2	AC122571_2	Continuation (3 of	c 926	17	1.7	128168	10	AL806526	AL806526 Mouse DNA
c 854	17	1.7	110000	2	AC125066_4	Continuation (5 of	c 927	17	1.7	128228	9	AL138810	AL138810 Human DNA
855	17	1.7	110000	2	AL353584_4	Continuation (5 of	928	17	1.7	128228	9	HSA298105	AJ298105 Homo sapi
856	17	1.7	110000	2	AL672265_0	AL672265 Homo sapi	c 929	17	1.7	128425	2	AC128115	AC128115 Rattus no
857	17	1.7	110000	2	LMFLCHR18_04	Continuation (5 of	c 930	17	1.7	128493	2	AC109964	AC109964 Rattus no
858	17	1.7	110000	2	LMFLCHR31_00	AL499621 Leishmani	931	17	1.7	128600	9	AC005192	AC005192 Homo sapi
859	17	1.7	110000	2	LMFLCHR32_11	Continuation (12 o	932	17	1.7	128815	2	AL773559	AL773559 Sus scrof
860	17	1.7	110000	2	AC055726_3	Continuation (4 of	c 933	17	1.7	128915	2	AC126000	AC126000 Rattus no
861	17	1.7	110000	9	HSY237C10_2	Continuation (3 of	c 934	17	1.7	128968	9	AL356858	AL356858 Human DNA
c 862	17	1.7	110898	9	AL355973	AL355973 Human DNA	c 935	17	1.7	129096	2	AC026412	AC026412 Homo sapi
c 863	17	1.7	111075	2	AC114337	AC114337 Homo sapi	c 936	17	1.7	129992	9	AL354897	AL354897 Human DNA
c 864	17	1.7	111104	9	AC117389	AC117389 Homo sapi	c 937	17	1.7	130474	2	AC121466	AC121466 Rattus no
c 865	17	1.7	111187	2	AP001801	AP001801 Homo sapi	c 938	17	1.7	130754	2	AC008508	AC008508 Homo sapi
866	17	1.7	111234	9	AC013451	AC013451 Homo sapi	939	17	1.7	131546	9	HS142O9	AL035687 Human DNA
867	17	1.7	111762	9	AC103851	AC103851 Homo sapi	940	17	1.7	131637	9	CNS01DUC	AL133239 Human chr
868	17	1.7	112180	9	HSJ322G13	AL096677 Human DNA	941	17	1.7	132035	2	AC125548	AC125548 Rattus no



942	17	1.7	132147	2	AC125826	AC125826	Rattus no
943	17	1.7	132171	3	AC008370	AC008370	Drosophil
944	17	1.7	132614	10	AC021756	AC021756	Mus muscu
945	17	1.7	132710	9	AC093535	AC093535	Homo sapi
946	17	1.7	132878	2	AC108187	AC108187	Felis cat
947	17	1.7	132953	9	AC024221	AC024221	Homo sapi
948	17	1.7	133294	2	AC118310	AC118310	Rattus no
949	17	1.7	133418	9	AL353709	AL353709	Human DNA
950	17	1.7	133477	10	AL646091	AL646091	Mouse DNA
951	17	1.7	133730	2	AC104294	AC104294	Rattus no
952	17	1.7	133789	9	AL442126	AL442126	Human DNA
953	17	1.7	133790	9	AC010280	AC010280	Homo sapi
954	17	1.7	134019	3	AC006471	AC006471	Drosophil
955	17	1.7	134250	9	AL139044	AL139044	Human DNA
956	17	1.7	134506	9	HS357116	HS357116	Homo sapi
957	17	1.7	134518	2	OSJN00134	OSJN00134	Oryza sat
958	17	1.7	135062	9	AC026691	AC026691	Homo sapi
959	17	1.7	135584	2	AC125496	AC125496	Oryza sat
960	17	1.7	136093	2	AC067894	AC067894	Homo sapi
961	17	1.7	136219	9	AP000697	AP000697	Homo sapi
962	17	1.7	137217	2	HSA2111L10	HSA2111L10	Homo sapi
963	17	1.7	137586	9	AC106775	AC106775	Homo sapi
964	17	1.7	137740	1	D90900	D90900	Synechocyst
965	17	1.7	137860	9	AC104826	AC104826	Homo sapi
966	17	1.7	138616	2	AC115525	AC115525	Rattus no
967	17	1.7	138982	2	AC104550	AC104550	Mus muscu
968	17	1.7	139218	2	AC019289	AC019289	Homo sapi
969	17	1.7	139290	2	AC068096	AC068096	Homo sapi
970	17	1.7	139924	2	AC116512	AC116512	Mus muscu
971	17	1.7	140774	2	AC108904	AC108904	Felis cat
972	17	1.7	140949	2	AC117339	AC117339	Rattus no
973	17	1.7	141120	9	AC002528	AC002528	Human BAC
974	17	1.7	141670	9	AC104170	AC104170	Homo sapi
975	17	1.7	141815	9	AC112716	AC112716	Homo sapi
976	17	1.7	142017	2	HSA196H18	HSA196H18	Homo sapi
977	17	1.7	142085	2	AC036153	AC036153	Homo sapi
978	17	1.7	142143	2	AC129352	AC129352	Homo sapi
979	17	1.7	142195	2	AC010669	AC010669	Drosophil
980	17	1.7	142388	2	AC128446	AC128446	Rattus no
981	17	1.7	142478	2	AC106966	AC106966	Rattus no
982	17	1.7	142837	2	AC113232	AC113232	Canis fam
983	17	1.7	142955	8	AP003349	AP003349	Oryza sat
984	17	1.7	143062	2	AC109092	AC109092	Rattus no
985	17	1.7	143264	2	AC119547	AC119547	Rattus no
986	17	1.7	143418	2	AC130228	AC130228	Rattus no
987	17	1.7	143534	2	AC036128	AC036128	Homo sapi
988	17	1.7	143611	9	AC068992	AC068992	Homo sapi
989	17	1.7	143676	2	AC026435	AC026435	Homo sapi
990	17	1.7	143684	2	AC107323	AC107323	Felis cat
991	17	1.7	143879	2	AC130724	AC130724	Oryza sat
992	17	1.7	144000	9	AC007621	AC007621	Homo sapi
993	17	1.7	144328	9	AC009695	AC009695	Homo sapi
994	17	1.7	144514	2	AC115534	AC115534	Sus scrof
995	17	1.7	144609	9	AC021269	AC021269	Homo sapi
996	17	1.7	144675	2	AC116060	AC116060	Rattus no
997	17	1.7	144724	8	AC078948	AC078948	Oryza sat
998	17	1.7	144967	2	AC008963	AC008963	Homo sapi
999	17	1.7	145216	3	AC092224	AC092224	Drosophil
1000	17	1.7	145265	8	AP003418	AP003418	Oryza sat

ALIGNMENTS

RESULT 1  
AC113646/c  
LOCUS AC113646 151778 bp DNA linear HTG 13-JUL-2002  
DEFINITION Rattus norvegicus clone CH230-324G7, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 35 unordered pieces.  
ACCESSION AC113646  
VERSION AC113646.3 GI:21738663  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE Norway rat.

ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 151778)  
Muzny D.M., Adams C., Adio-Oduola B., Ali-osman, F.R., Allen C.,  
Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,  
Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,  
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,  
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlssoon, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,  
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G.,  
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,  
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

TITLE  
JOURNAL

Unpublished  
2 (bases 1 to 151778)  
Worley, K.C.

REFERENCE  
AUTHORS

Direct Submission  
Submitted (05-MAR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

TITLE  
JOURNAL

3 (bases 1 to 151778)  
Worley, K.C.  
Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Département  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS

On Jul 12, 2002 this sequence version replaced gi:19525805.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GRFA  
Center clone name: CH230-324G7  
----- Summary Statistics  
Sequencing vector: Plasmid:  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 123655 bases at least Q40

COMMENT

Consensus quality: 126566 bases at least Q30  
Consensus quality: 129561 bases at least Q20

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
NOTE: This is a 'working draft' sequence. It currently consists of 35 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1002: contig of 1002 bp in length  
1003 1102: gap of unknown length  
1103 2439: contig of 1337 bp in length  
2440 2539: gap of unknown length  
2540 3615: contig of 1076 bp in length  
3616 3715: gap of unknown length  
3716 5109: contig of 1394 bp in length  
5110 5209: gap of unknown length  
5210 6650: contig of 1441 bp in length  
6651 6750: gap of unknown length  
6751 8244: contig of 1494 bp in length  
8245 8344: gap of unknown length  
8345 9594: contig of 1250 bp in length  
9595 9694: gap of unknown length  
9695 11425: contig of 1731 bp in length  
11426 11525: gap of unknown length  
11526 13702: contig of 2177 bp in length  
13703 13802: gap of unknown length  
13803 16415: contig of 2613 bp in length  
16416 16515: gap of unknown length  
16516 18030: contig of 1515 bp in length  
18031 18130: gap of unknown length  
18131 20246: contig of 2116 bp in length  
20247 20346: gap of unknown length  
20347 22138: contig of 1792 bp in length  
22139 22238: gap of unknown length  
22239 24024: contig of 1786 bp in length  
24025 24124: gap of unknown length  
24125 27113: contig of 2989 bp in length  
27114 27213: gap of unknown length  
27214 30060: contig of 2847 bp in length  
30061 30160: gap of unknown length  
30161 33562: contig of 3402 bp in length  
33563 33662: gap of unknown length  
33663 35820: contig of 2158 bp in length  
35821 35920: gap of unknown length  
35921 38286: contig of 2366 bp in length  
38287 38386: gap of unknown length  
38387 41808: contig of 3422 bp in length  
41809 41908: gap of unknown length  
41909 45894: contig of 3986 bp in length  
45895 45994: gap of unknown length  
45995 48647: contig of 2653 bp in length  
48648 48747: gap of unknown length  
48748 54666: contig of 5919 bp in length  
54667 54766: gap of unknown length  
54767 59394: contig of 4628 bp in length  
59395 59494: gap of unknown length  
59495 65824: contig of 6330 bp in length  
65825 65924: gap of unknown length  
65925 72889: contig of 6965 bp in length  
72890 72989: gap of unknown length  
72990 80055: contig of 7066 bp in length  
80056 80155: gap of unknown length  
80156 86569: contig of 6414 bp in length  
86570 86669: gap of unknown length  
86670 94191: contig of 7522 bp in length  
94192 94291: gap of unknown length  
94292 103725: contig of 9434 bp in length  
103726 103825: gap of unknown length

\* 103826 110175: contig of 6350 bp in length  
\* 110176 110275: gap of unknown length  
\* 110276 118450: contig of 8175 bp in length  
\* 118451 118550: gap of unknown length  
\* 118551 129815: contig of 11265 bp in length  
\* 129816 129915: gap of unknown length  
\* 129916 139145: contig of 9230 bp in length  
\* 139146 139245: gap of unknown length  
\* 139246 151778: contig of 12533 bp in length.

FEATURES  
source  
1. 151778  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-324G7"

BASE COUNT 43201 a 29596 c 29588 g 45822 t 3571 others  
ORIGIN

Query Match 2.4%; Score 23; DB 2; Length 151778;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 845 TCCTACCAAACTCAATAAAAGT 867  
|||||  
Db 32134 TCCTACCAAACTCAATAAAAGT 32112

RESULT 2  
AC099573  
LOCUS  
DEFINITION Mus musculus clone RP23-28E6, WORKING DRAFT SEQUENCE, 5 unordered pieces.  
AC099573  
AC099573.2 GI:22325167  
VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM  
REFERENCE 1 (bases 1 to 214265)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triggilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (16-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 214265)  
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (20-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 20, 2002 this sequence version replaced gi:16946021.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: L13795  
Center clone name: 28\_E6

----- Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 213019 bases at least Q40  
Consensus quality: 213493 bases at least Q30  
Consensus quality: 213665 bases at least Q20  
Insert size: 210000; agarose-fp  
Insert size: 213865; sum-of-contigs  
Quality coverage: 14.5 in Q20 bases; agarose-fp  
Quality coverage: 14.3 in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 16219: contig of 16219 bp in length  
\* 16220 16319: gap of 100 bp  
\* 16320 120317: contig of 103998 bp in length  
\* 120318 120417: gap of 100 bp  
\* 120418 163351: contig of 42934 bp in length  
\* 163352 163451: gap of 100 bp  
\* 163452 210716: contig of 47265 bp in length  
\* 210717 210816: gap of 100 bp  
\* 210817 214265: contig of 3449 bp in length.

FEATURES  
source  
Location/Qualifiers  
1. .214265  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-28E6"  
/clone\_lib="RPCI-23 Female Mouse BAC"

misc\_feature  
1. .16219  
/note="assembly\_fragment  
clone\_end:SP6  
vector\_side:left"  
16320. .120317  
/note="assembly\_fragment"  
120418. .163351  
/note="assembly\_fragment"  
163452. .210716

misc\_feature  
/note="assembly\_fragment"  
210817. .214265  
/note="assembly\_fragment  
clone\_end:T7  
vector\_side:right"

BASE COUNT 72086 a 41630 c 41065 g 59084 t 400 others  
ORIGIN

Query Match 2.3%; Score 22; DB 2; Length 214265;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 GTGGATTAAACCAAAAAATGG 287  
|||||  
Db 160881 GTGGATTAAACCAAAAAATGG 160902

RESULT 3  
AC114589/c

LOCUS AC114589 62004 bp DNA linear HTG 10-MAR-2002  
DEFINITION Mus musculus clone RP23-365013, LOW-PASS SEQUENCE SAMPLING.  
ACCESSION AC114589  
VERSION AC114589.1 GI:19311117  
KEYWORDS HTG; HTGS PHASE0.  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 62004)  
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
TITLE Mus musculus, clone RP23-365013  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 62004)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: L23822  
Center clone name: 365\_O13

-----  
\* NOTE: This record contains 77 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads



\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 704: contig of 704 bp in length  
705 804: gap of 100 bp  
805 1480: contig of 676 bp in length  
1481 1580: gap of 100 bp  
1581 2282: contig of 702 bp in length  
2283 2382: gap of 100 bp  
2383 3090: contig of 708 bp in length  
3091 3190: gap of 100 bp  
3191 3846: contig of 656 bp in length  
3847 3946: gap of 100 bp  
3947 4663: contig of 717 bp in length  
4664 4763: gap of 100 bp  
4764 5456: contig of 693 bp in length  
5457 5556: gap of 100 bp  
5557 6248: contig of 692 bp in length  
6249 6348: gap of 100 bp  
6349 7061: contig of 713 bp in length  
7062 7161: gap of 100 bp  
7162 7865: contig of 704 bp in length  
7866 7965: gap of 100 bp  
7966 8679: contig of 714 bp in length  
8680 8779: gap of 100 bp  
8780 9484: contig of 705 bp in length  
9485 9584: gap of 100 bp  
9585 10300: contig of 716 bp in length  
10301 10400: gap of 100 bp  
10401 11109: contig of 709 bp in length  
11110 11209: gap of 100 bp  
11210 11914: contig of 705 bp in length  
11915 12014: gap of 100 bp  
12015 12713: contig of 699 bp in length  
12714 12813: gap of 100 bp  
12814 13523: contig of 710 bp in length  
13524 13623: gap of 100 bp  
13624 14333: contig of 710 bp in length  
14334 14433: gap of 100 bp  
14434 15147: contig of 714 bp in length  
15148 15247: gap of 100 bp  
15248 15951: contig of 704 bp in length  
15952 16051: gap of 100 bp  
16052 16774: contig of 723 bp in length  
16775 16874: gap of 100 bp  
16875 17579: contig of 705 bp in length  
17580 17679: gap of 100 bp  
17680 18398: contig of 719 bp in length  
18399 18498: gap of 100 bp  
18499 19212: contig of 714 bp in length  
19213 19312: gap of 100 bp  
19313 20021: contig of 709 bp in length  
20022 20121: gap of 100 bp  
20122 20834: contig of 713 bp in length  
20835 20934: gap of 100 bp  
20935 21645: contig of 711 bp in length  
21646 21745: gap of 100 bp  
21746 22454: contig of 709 bp in length  
22455 22554: gap of 100 bp  
22555 23257: contig of 703 bp in length  
23258 23357: gap of 100 bp  
23358 24008: contig of 651 bp in length  
24009 24108: gap of 100 bp  
24109 24812: contig of 704 bp in length  
24813 24912: gap of 100 bp  
24913 25625: contig of 713 bp in length  
25626 25725: gap of 100 bp  
25726 26449: contig of 724 bp in length

26450 26549: gap of 100 bp  
26550 27257: contig of 708 bp in length  
27258 27357: gap of 100 bp  
27358 28069: contig of 712 bp in length  
28070 28169: gap of 100 bp  
28170 28864: contig of 695 bp in length  
28865 28964: gap of 100 bp  
28965 29675: contig of 711 bp in length  
29676 29775: gap of 100 bp  
29776 30477: contig of 702 bp in length  
30478 30577: gap of 100 bp  
30578 31292: contig of 715 bp in length  
31293 31392: gap of 100 bp  
31393 32112: contig of 720 bp in length  
32113 32212: gap of 100 bp  
32213 32919: contig of 707 bp in length  
32920 33019: gap of 100 bp  
33020 33735: contig of 716 bp in length  
33736 33835: gap of 100 bp  
33836 34547: contig of 712 bp in length  
34548 34647: gap of 100 bp  
34648 35354: contig of 707 bp in length  
35355 35454: gap of 100 bp  
35455 36127: contig of 673 bp in length  
36128 36227: gap of 100 bp  
36228 36940: contig of 713 bp in length  
36941 37040: gap of 100 bp  
37041 37751: contig of 711 bp in length  
37752 37851: gap of 100 bp  
37852 38545: contig of 694 bp in length  
38546 38645: gap of 100 bp  
38646 39357: contig of 712 bp in length  
39358 39457: gap of 100 bp  
39458 40157: contig of 700 bp in length  
40158 40257: gap of 100 bp  
40258 40972: contig of 715 bp in length  
40973 41072: gap of 100 bp  
41073 41791: contig of 719 bp in length  
41792 41891: gap of 100 bp  
41892 42599: contig of 708 bp in length  
42600 42699: gap of 100 bp  
42700 43402: contig of 703 bp in length  
43403 43502: gap of 100 bp  
43503 44211: contig of 709 bp in length  
44212 44311: gap of 100 bp  
44312 45024: contig of 713 bp in length  
45025 45124: gap of 100 bp  
45125 45835: contig of 711 bp in length  
45836 45935: gap of 100 bp  
45936 46639: contig of 704 bp in length  
46640 46739: gap of 100 bp  
46740 47452: contig of 713 bp in length  
47453 47552: gap of 100 bp  
47553 48255: contig of 703 bp in length  
48256 48355: gap of 100 bp  
48356 49078: contig of 723 bp in length  
49079 49178: gap of 100 bp  
49179 49899: contig of 721 bp in length  
49900 49999: gap of 100 bp  
50000 50716: contig of 717 bp in length  
50717 50816: gap of 100 bp  
50817 51528: contig of 712 bp in length  
51529 51628: gap of 100 bp  
51629 52338: contig of 710 bp in length  
52339 52438: gap of 100 bp  
52439 53146: contig of 708 bp in length  
53147 53246: gap of 100 bp  
53247 53959: contig of 713 bp in length  
53960 54059: gap of 100 bp  
54060 54762: contig of 703 bp in length  
54763 54862: gap of 100 bp  
54863 55547: contig of 685 bp in length  
55548 55647: gap of 100 bp

Query Match 2.2%; Score 21; DB 2; Length 62004;  
Best Local Similarity 100.0%; Pred. No. 6.5;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 626 AAAAAGCCAACCATATTG 646  
Db 56146 AAAAAGCCAACCATATTG 56126

RESULT 4  
AL355985/c 64969 bp DNA linear PRI 15-NOV-2001  
LOCUS Human DNA sequence from clone RP11-98017 on chromosome 9, complete  
DEFINITION sequence.  
ACCESSION AL355985  
VERSION AL355985.18 GI:16972867  
KEYWORDS HTG.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 64969)  
AUTHORS Skuce,C.  
TITLE Direct Submission  
JOURNAL Submitted (15-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENT On Nov 16, 2001 this sequence version replaced gi:16416171.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em:, EMBL; Sw:,  
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP  
database can be found at

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping  
Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr9  
RP11-98017 is from the library RPCI-11.1 constructed by the group  
of Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone  
RP11-98017 It may be shorter because we sequence overlapping  
sections only once, except for a short overlap.  
The true left end of clone RP11-38824 is at 62970 in this  
sequence. The true right end of clone RP11-383M4 is at 2000 in this  
sequence.

FEATURES Location/Qualifiers  
source 1..64969  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="9"  
/clone="RP11-98017"  
/clone\_lib="RPCI-11.1"  
BASE COUNT 18132 a 13055 c 13036 g 20746 t  
ORIGIN

Query Match 2.2%; Score 21; DB 9; Length 64969;  
Best Local Similarity 100.0%; Pred. No. 6.5;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 898 GATGAGAAGGAATAGAAATTG 918  
Db 41352 GATGAGAAGGAATAGAAATTG 41332

RESULT 5  
AC087568 150199 bp DNA linear PRI 02-NOV-2001  
LOCUS Pan troglodytes clone RP43-135M11, complete sequence.  
DEFINITION AC087568  
ACCESSION AC087568  
VERSION AC087568.2 GI:16596569  
KEYWORDS HTG.  
SOURCE Pan troglodytes.

ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
REFERENCE 1 (bases 1 to 150199)  
AUTHORS Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,  
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,  
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,  
Ho,S.-L., Idol,J.R., Karlins,E., Laric,P., Lee-Lin,S.-Q.,  
Legaspi,R., Maduro,Q.L., Maduro,V.B., Masiello,C., Mastrian,S.D.,  
McCloskey,J.C., McDowell,J., Pearson,R., Prasad,A., Shevchenko,Y.,  
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W.,  
Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,  
Zhang,L.-H. and Green,E.D.

NISC Comparative Sequencing Initiative  
Unpublished  
REFERENCE 2 (bases 1 to 150199)  
AUTHORS Green,E.D.  
TITLE Direct Submission  
JOURNAL Submitted (10-JAN-2001) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA  
REFERENCE 3 (bases 1 to 150199)  
AUTHORS Green,E.D.  
TITLE Direct Submission  
JOURNAL Submitted (02-NOV-2001) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA  
On Nov 2, 2001 this sequence version replaced gi:12061434.

COMMENT ----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: http://www.nisc.nih.gov  
Contact: nisc.mouse@nhgri.nih.gov  
----- Project Information  
Center project name: aof  
Center clone name: 135M11

This sequence was finished as follows unless otherwise noted:  
all regions were double-stranded, sequenced with an  
alternate chemistry, or covered by high quality data  
(i.e., phred quality >= 30); an attempt was made to resolve  
all sequencing problems, such as compressions and repeats;  
all regions were covered by at least one plasmid subclone  
or more than one M13 subclone; and the assembly was confirmed  
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of  
this clone unless otherwise noted. If there are overlapping  
clones, the overlaps are noted in the beginning and end of  
the Features section.  
Location/Qualifiers  
1..150199  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="RP43-135M11"  
/clone\_lib="RP43"  
BASE COUNT 36501 a 36322 c 37037 g 40339 t  
ORIGIN

Query Match 2.2%; Score 21; DB 9; Length 150199;  
Best Local Similarity 100.0%; Pred. No. 6.5;

Best Local Similarity 100.0%; Pred. NO. 6:3; Mismatches 0; Indels 0; Gaps 0; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 ATGTTTGGACCTATATTAGCT 421  
|||||  
Db 79411 ATGTTTGGACCTATATTAGCT 79431

RESULT 6  
AC1111597  
LOCUS  
DEFINITION Rattus norvegicus clone CH230-150P11, linear HTG 13-JUL-2002  
\*\*\*, 62 unordered pieces.  
AC1111597 169574 bp DNA  
VERSION AC111597.2 GI:21736155  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
1 (bases 1 to 169574)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,  
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 169574)  
Worley,K.C.  
Direct Submission  
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 169574)  
Worley,K.C.  
Direct Submission

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE

JOURNAL Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
COMMENT On Jul 12, 2002 this sequence version replaced gi:18701364.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GMQL  
Center clone name: CH230-150P11  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 113501 bases at least Q40  
Consensus quality: 119212 bases at least Q30  
Consensus quality: 124054 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 62 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1152: contig of 1152 bp in length  
\* 1153 1252: gap of unknown length  
\* 1253 2646: contig of 1394 bp in length  
\* 2647 2746: gap of unknown length  
\* 2747 4284: contig of 1538 bp in length  
\* 4285 4384: gap of unknown length  
\* 4385 5430: contig of 1046 bp in length  
\* 5431 5530: gap of unknown length  
\* 5531 6682: contig of 1152 bp in length  
\* 6683 6782: gap of unknown length  
\* 6783 8328: contig of 1546 bp in length  
\* 8329 8428: gap of unknown length  
\* 8429 9654: contig of 1226 bp in length  
\* 9655 9754: gap of unknown length  
\* 9755 11055: contig of 1301 bp in length  
\* 11056 11155: gap of unknown length  
\* 11156 12280: contig of 1125 bp in length  
\* 12281 12380: gap of unknown length  
\* 12381 13455: contig of 1075 bp in length  
\* 13456 13555: gap of unknown length  
\* 13556 15071: contig of 1516 bp in length  
\* 15072 15171: gap of unknown length  
\* 15172 16403: contig of 1232 bp in length  
\* 16404 16503: gap of unknown length  
\* 16504 17701: contig of 1198 bp in length  
\* 17702 17801: gap of unknown length  
\* 17802 19484: contig of 1683 bp in length  
\* 19485 19584: gap of unknown length  
\* 19585 21172: contig of 1588 bp in length  
\* 21173 21272: gap of unknown length  
\* 21273 22768: contig of 1496 bp in length  
\* 22769 22868: gap of unknown length  
\* 22869 24002: contig of 1134 bp in length  
\* 24003 24102: gap of unknown length  
\* 24103 25537: contig of 1435 bp in length  
\* 25538 25637: gap of unknown length  
\* 25638 26903: contig of 1266 bp in length  
\* 26904 27003: gap of unknown length  
\* 27004 28985: contig of 1982 bp in length  
\* 28986 29085: gap of unknown length  
\* 29086 31277: contig of 2192 bp in length  
\* 31278 31377: gap of unknown length  
\* 31378 33368: contig of 1991 bp in length



* 33369	* 33468: gap of	unknown length
* 33469	* 34518: contig of	1050 bp in length
* 34519	* 34618: gap of	unknown length
* 34619	* 36488: contig of	1870 bp in length
* 36489	* 36588: gap of	unknown length
* 36589	* 38145: contig of	1557 bp in length
* 38146	* 38245: gap of	unknown length
* 38246	* 39774: contig of	1529 bp in length
* 39775	* 39874: gap of	unknown length
* 39875	* 41463: contig of	1589 bp in length
* 41464	* 41563: gap of	unknown length
* 41564	* 42961: contig of	1398 bp in length
* 42962	* 43061: gap of	unknown length
* 43062	* 45065: contig of	2004 bp in length
* 45066	* 45165: gap of	unknown length
* 45166	* 47102: contig of	1937 bp in length
* 47103	* 47202: gap of	unknown length
* 47203	* 49519: contig of	2317 bp in length
* 49520	* 49619: gap of	unknown length
* 49620	* 51310: contig of	1691 bp in length
* 51311	* 51410: gap of	unknown length
* 51411	* 52960: contig of	1550 bp in length
* 52961	* 53060: gap of	unknown length
* 53061	* 55201: contig of	2141 bp in length
* 55202	* 55301: gap of	unknown length
* 55302	* 57573: contig of	2272 bp in length
* 57574	* 57673: gap of	unknown length
* 57674	* 60208: contig of	2535 bp in length
* 60209	* 60308: gap of	unknown length
* 60309	* 62351: contig of	2043 bp in length
* 62352	* 62451: gap of	unknown length
* 62452	* 64851: contig of	2400 bp in length
* 64852	* 64951: gap of	unknown length
* 64952	* 67104: contig of	2153 bp in length
* 67105	* 67204: gap of	unknown length
* 67205	* 69433: contig of	2229 bp in length
* 69434	* 69533: gap of	unknown length
* 69534	* 72226: contig of	2693 bp in length
* 72227	* 72326: gap of	unknown length
* 72327	* 76213: contig of	3887 bp in length
* 76214	* 76313: gap of	unknown length
* 76314	* 78709: contig of	2396 bp in length
* 78710	* 78809: gap of	unknown length
* 78810	* 81500: contig of	2691 bp in length
* 81501	* 81600: gap of	unknown length
* 81601	* 84321: contig of	2721 bp in length
* 84322	* 84421: gap of	unknown length
* 84422	* 87220: contig of	2799 bp in length
* 87221	* 87320: gap of	unknown length
* 87321	* 92378: contig of	5058 bp in length
* 92379	* 92478: gap of	unknown length
* 92479	* 95106: contig of	2628 bp in length
* 95107	* 95206: gap of	unknown length
* 95207	* 99016: contig of	3810 bp in length
* 99017	* 99116: gap of	unknown length
* 99117	* 101692: contig of	2576 bp in length
* 101693	* 101792: gap of	unknown length
* 101793	* 104836: contig of	3044 bp in length
* 104837	* 104936: gap of	unknown length
* 104937	* 110703: contig of	5767 bp in length
* 110704	* 110803: gap of	unknown length
* 110804	* 113760: contig of	2957 bp in length
* 113761	* 113860: gap of	unknown length
* 113861	* 118784: contig of	4924 bp in length

Query Match 2.2%; Score 21; DB 2; Length 169574;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 AAGACAAAGCAGAGGGTGAAA 172  
|||||

Db 3741 AAGACAAAGCAGAGGGTGAAA 3761  
|||||

RESULT 7  
AC091001  
LOCUS AC091001 177066 bp DNA linear PRI 06-DEC-2001  
DEFINITION Papio cynocephalus anubis clone RP41-36M16, complete sequence.  
ACCESSION AC091001  
VERSION AC091001.2 GI:17386264  
KEYWORDS HTG.  
SOURCE  
ORGANISM Papio cynocephalus anubis.  
Papio cynocephalus anubis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
Cercopithecinae; Papio.  
REFERENCE 1 (bases 1 to 177066)  
AUTHORS Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,  
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,  
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,  
Ho,S.-L., Idol,J.R., Karlins,E., Laric,P., Lee-Lin,S.-Q.,  
Legaspi,R., Maduro,Q.L., Maduro,V.B., Masiello,C., Mastrian,S.D.,  
McCloskey,J.C., McDowell,J., Pearson,R., Prasad,A., Shevchenko,Y.,  
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W.,  
Tsurgon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,  
Zhang,L.-H. and Green,E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
REFERENCE 2 (bases 1 to 177066)  
AUTHORS Green,E.D.  
TITLE Direct Submission  
JOURNAL Submitted (26-MAR-2001) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA  
REFERENCE 3 (bases 1 to 177066)  
AUTHORS Green,E.D.  
TITLE Direct Submission  
JOURNAL Submitted (06-DEC-2001) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA  
On Dec 6, 2001 this sequence version replaced gi:13447439.  
COMMENT  
----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: http://www.nisc.nih.gov  
Contact: nisc.mouse@nhgri.nih.gov  
----- Project Information  
Center project name: asg  
Center clone name: 036M16

This sequence was finished as follows unless otherwise noted:  
all regions were double-stranded, sequenced with an  
alternate chemistry, or covered by high quality data  
(i.e., phred quality >= 30); an attempt was made to resolve  
all sequencing problems, such as compressions and repeats;  
all regions were covered by at least one plasmid subclone  
or more than one M13 subclone; and the assembly was confirmed  
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of  
this clone unless otherwise noted. If there are overlapping  
clones, the overlaps are noted in the beginning and end of  
the Features section.

FEATURES  
source Location/Qualifiers  
1. 177066  
/organism="Papio cynocephalus anubis"  
/db\_xref="taxon:9555"  
/clone="RP41-36M16"  
/clone\_lib="RP41"  
misc\_feature 2718..3002  
/note="single clone coverage"  
misc\_feature 3862..4236  
/note="single clone coverage"  
misc\_feature 82768..82855  
/note="single clone coverage"  
misc\_feature 83766..83796  
/note="single clone coverage"  
misc\_feature 92916..92927

misc\_feature /note="single clone coverage"  
93063. .93191  
misc\_feature /note="single clone coverage"  
94176. .94491  
misc\_feature /note="single clone coverage"  
142486. .142509  
BASE COUNT 42959 a 42842 c 44880 g 46385 t  
ORIGIN

Query Match 2.2%; Score 21; DB 9; Length 177066;  
Best Local Similarity 100.0%; Pred. NO. 6.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 ATGTTTGGACCTATATTAGCT 421  
|||||  
Db 100974 ATGTTTGGACCTATATTAGCT 100994  
|||||

RESULT 8  
AC005089 AC005089 177893 bp DNA linear PRI 30-MAR-2002  
LOCUS Homo sapiens BAC clone CTA-315H11 from 7q11.2, complete sequence.  
DEFINITION AC005089  
AC005089  
VERSION AC005089.3 GI:19747251  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 177893)  
Sulston, J.E. and Waterston, R.  
Toward a complete human genome sequence  
Genome Res. 8 (11), 1097-1108 (1998)  
99063792  
9847074  
2 (bases 1 to 177893)  
Graves, T., Maas, J. and Hawkins, M.  
The sequence of Homo sapiens BAC clone CTA-315H11  
Unpublished (2001)  
3 (bases 1 to 177893)  
Waterston, R.H.  
Direct Submission  
Submitted (12-JUN-1998) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
4 (bases 1 to 177893)  
Waterston, R.H.  
Direct Submission  
Submitted (13-AUG-1999) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
5 (bases 1 to 177893)  
Waterston, R.  
Direct Submission  
Submitted (02-OCT-2000) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
6 (bases 1 to 177893)  
Waterston, R.H.  
Direct Submission  
Submitted (27-MAR-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
7 (bases 1 to 177893)  
Waterston, R.  
Direct Submission  
Submitted (30-MAR-2002) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Mar 27, 2002 this sequence version replaced gi:5732140.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc

Contact: sapiens@watson.wustl.edu  
----- Summary Statistics  
Center project name: H\_RG315H11  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see  
http://www.nhgri.nih.gov/DIR/OTB/CHR7, send  
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:  
Clone CTA-315H11 is from a release of the human BAC library C1TB-HS-A. The library contains cloned DNA from human sperm. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). The clone is available from Research Genetics, Inc. (http://www.resgen.com).  
VECTOR: pBelobAC11  
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the right is RG208H19. Actual start of this clone is at base position 1 of RG315H11; actual end is at base position 177893 of RG315H11.  
Location/Qualifiers  
1. .177893  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7q11.2"  
/clone="CTA-315H11"  
/clone\_lib="C1TB-HS-A"  
51. .193  
/rpt\_family="MIR"  
194. .282  
/rpt\_family="MIR"  
613. .910  
/rpt\_family="Alu"  
948. .1044  
/rpt\_family="MIR"  
1016. .1515  
/note="match to EST BF840093 (NID:g12192535)"  
1260. .1317  
/rpt\_family="GA-rich"  
1431. .1733  
/rpt\_family="Alu"  
1743. .1867  
/rpt\_family="Alu"  
1871. .2170  
/rpt\_family="Alu"  
2324. .2405  
/rpt\_family="MIR"  
2406. .2675  
/rpt\_family="Alu"  
2863. .3142  
/rpt\_family="ERV1"

repeat\_region  
repeat\_region  
repeat\_region  
repeat\_region  
misc\_feature  
repeat\_region  
repeat\_region  
repeat\_region  
repeat\_region  
repeat\_region  
repeat\_region  
repeat\_region

repeat\_region 3146. .3283 /rpt\_family="Alu"  
repeat\_region 3285. .3571 /rpt\_family="Alu"  
repeat\_region 3577. .3597 /rpt\_family="TAAA)n"  
repeat\_region 3598. .4006 /rpt\_family="ERV1"  
repeat\_region 4007. .4307 /rpt\_family="Alu"  
repeat\_region 4308. .4353 /rpt\_family="ERV1"  
repeat\_region 4526. .4618 /rpt\_family="MIR"  
repeat\_region 4694. .4920 /rpt\_family="Alu"  
repeat\_region 5134. .5445 /rpt\_family="Alu"  
repeat\_region 5511. .5817 /rpt\_family="Alu"  
repeat\_region 5829. .6163 /rpt\_family="Alu"  
repeat\_region 6252. .6315 /rpt\_family="MIR"  
repeat\_region 6316. .6442 /rpt\_family="Alu"  
repeat\_region 6443. .6464 /rpt\_family="MIR"  
repeat\_region 6470. .6672 /rpt\_family="L1"  
misc\_feature 6569. .7029 /note="match to EST BE784135 (NID:g10205333)"  
repeat\_region 7032. .7298 /rpt\_family="Alu"  
repeat\_region 7304. .7326 /rpt\_family="(CAAA)n"  
repeat\_region 7359. .7639 /rpt\_family="Alu"  
repeat\_region 7648. .7748 /rpt\_family="MIR"  
repeat\_region 8239. .8293 /rpt\_family="MIR"  
repeat\_region 8409. .8438 /rpt\_family="MER1\_type"  
repeat\_region 8439. .8744 /rpt\_family="Alu"  
repeat\_region 8745. .8857 /rpt\_family="MER1\_type"  
repeat\_region 8858. .9160 /rpt\_family="Alu"  
misc\_feature 9150. .9410 /note="match to EST BE784135 (NID:g10205333)"  
repeat\_region 9161. .9284 /rpt\_family="MER1\_type"  
repeat\_region 9441. .9515 /rpt\_family="L2"  
repeat\_region 9516. .9812 /rpt\_family="Alu"  
repeat\_region 9813. .10073 /rpt\_family="L2"  
repeat\_region 10076. .10588 /rpt\_family="(TGGA)n"  
repeat\_region 10614. .10655 /rpt\_family="polypyrimidine"  
repeat\_region 10815. .11134 /rpt\_family="Alu"  
repeat\_region 11216. .11245 /rpt\_family="AT\_rich"  
repeat\_region 11246. .11537 /rpt\_family="Alu"  
repeat\_region 11556. .11690 /rpt\_family="Alu"  
repeat\_region 11807. .11891

repeat\_region /rpt\_family="MIR"  
12186. .12321  
Query Match 2.2% Score 21; DB 9; Length 177893;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 401 ATGTTTGGACCTATATTAGCT 421  
Db 96884 ATGTTTGGACCTATATTAGCT 96904  
RESULT 9  
AC128082 178228 bp DNA linear HTG 19-JUL-2002  
LOCUS Rattus norvegicus clone CH230-362C16, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION \*\*\*, 64 unordered pieces.  
ACCESSION AC128082 GI:21908679  
VERSION AC128082.1 HTG; HTGS\_PHASE1.  
KEYWORDS Rattus norvegicus.  
SOURCE Rattus norvegicus  
ORGANISM Rattus.  
REFERENCE 1 (bases 1 to 178228)  
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,K., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.  
Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 178228)  
AUTHORS Worley,K.C.  
TITLE Direct Submission  
JOURNAL Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One



COMMENT

Baylor Plaza, Houston, TX 77030, USA  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GZQW  
Center clone name: CH230-362C16  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 116771 bases at least Q40  
Consensus quality: 123909 bases at least Q30  
Consensus quality: 129462 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 64 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1173: contig of 1173 bp in length  
\* 1174 1273: gap of unknown length  
\* 1274 2435: contig of 1162 bp in length  
\* 2436 2535: gap of unknown length  
\* 2536 3567: contig of 1032 bp in length  
\* 3568 3667: gap of unknown length  
\* 3668 5230: contig of 1563 bp in length  
\* 5231 5330: gap of unknown length  
\* 5331 6732: contig of 1402 bp in length  
\* 6733 7990: gap of unknown length  
\* 7991 8090: contig of 1158 bp in length  
\* 8091 9641: gap of unknown length  
\* 9642 9741: gap of unknown length  
\* 9742 10776: contig of 1035 bp in length  
\* 10777 10876: gap of unknown length  
\* 10877 11905: contig of 1029 bp in length  
\* 11906 12005: gap of unknown length  
\* 12006 13320: contig of 1315 bp in length  
\* 13321 13420: gap of unknown length  
\* 13421 14678: contig of 1258 bp in length  
\* 14679 14778: gap of unknown length  
\* 14779 15851: contig of 1073 bp in length  
\* 15852 15951: gap of unknown length  
\* 15952 17593: contig of 1642 bp in length  
\* 17594 17693: gap of unknown length  
\* 17694 18911: contig of 1218 bp in length  
\* 18912 19011: gap of unknown length  
\* 19012 20145: contig of 1134 bp in length  
\* 20146 20245: gap of unknown length  
\* 20246 21556: contig of 1311 bp in length  
\* 21557 21656: gap of unknown length  
\* 21657 22715: contig of 1059 bp in length  
\* 22716 22815: gap of unknown length  
\* 22816 23949: contig of 1134 bp in length  
\* 23950 24049: gap of unknown length  
\* 24050 25331: contig of 1282 bp in length  
\* 25332 25431: gap of unknown length  
\* 25432 26617: contig of 1186 bp in length  
\* 26618 26717: gap of unknown length  
\* 26718 28435: contig of 1718 bp in length  
\* 28436 28535: gap of unknown length  
\* 28536 30660: contig of 2125 bp in length  
\* 30661 30760: gap of unknown length  
\* 30761 32263: contig of 1503 bp in length  
\* 32264 32363: gap of unknown length

32364 34165: contig of 1802 bp in length  
34166 34265: gap of unknown length  
34266 35302: contig of 1037 bp in length  
35303 35402: gap of unknown length  
35403 37282: contig of 1880 bp in length  
37283 37382: gap of unknown length  
37383 39160: contig of 1778 bp in length  
39161 39260: gap of unknown length  
39261 40880: contig of 1620 bp in length  
40881 40980: gap of unknown length  
40981 42436: contig of 1456 bp in length  
42437 42536: gap of unknown length  
42537 44381: contig of 1845 bp in length  
44382 44481: gap of unknown length  
44482 46391: contig of 1910 bp in length  
46392 46491: gap of unknown length  
46492 48409: contig of 1918 bp in length  
48410 48509: gap of unknown length  
48510 49931: contig of 1422 bp in length  
49932 50031: gap of unknown length  
50032 51371: contig of 1340 bp in length  
51372 51471: gap of unknown length  
51472 53362: contig of 1891 bp in length  
53363 53462: gap of unknown length  
53463 55170: contig of 1708 bp in length  
55171 55270: gap of unknown length  
55271 56697: contig of 1427 bp in length  
56698 56797: gap of unknown length  
56798 58037: contig of 1240 bp in length  
58038 58137: gap of unknown length  
58138 60361: contig of 2224 bp in length  
60362 60461: gap of unknown length  
60462 62335: contig of 1874 bp in length  
62336 62435: gap of unknown length  
62436 64906: contig of 2471 bp in length  
64907 65006: gap of unknown length  
65007 67475: contig of 2469 bp in length  
67476 67575: gap of unknown length  
67576 69653: contig of 2078 bp in length  
69654 73830: gap of unknown length  
73831 73930: contig of 4077 bp in length  
73931 77047: contig of 3117 bp in length  
77048 77147: gap of unknown length  
77148 80438: contig of 3291 bp in length  
80439 80538: gap of unknown length  
80539 83532: contig of 2994 bp in length  
83533 83632: gap of unknown length  
83633 85754: contig of 2122 bp in length  
85755 85854: gap of unknown length  
85855 89203: contig of 3349 bp in length  
89204 89303: gap of unknown length  
89304 92865: contig of 3562 bp in length  
92866 92965: gap of unknown length  
92966 95763: contig of 2798 bp in length  
95764 95863: gap of unknown length  
95864 99698: contig of 3835 bp in length  
99699 99798: gap of unknown length  
99799 104352: contig of 4554 bp in length  
104353 104452: gap of unknown length  
104453 108165: contig of 3713 bp in length  
108166 108265: gap of unknown length  
108266 111945: contig of 3680 bp in length  
111946 112045: gap of unknown length  
112046 116301: contig of 4256 bp in length  
116302 116401: gap of unknown length  
116402 120825: contig of 4424 bp in length  
120826 120925: gap of unknown length

Query Match 2.2%; Score 21; DB 2; Length 178228;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 AAGACAAAGCAGAGGCTGAAA 172  
Db 33628 AAGACAAAGCAGAGGCTGAAA 33648

RESULT 10  
AP003587  
LOCUS AP003587 343550 bp DNA linear BCT 28-NOV-2001  
DEFINITION Nostoc sp. PCC 7120 DNA, complete genome, section 7/19.  
ACCESSION AP003587 BA000019  
VERSION AP003587.1 GI:17130808  
KEYWORDS  
SOURCE Nostoc sp. PCC 7120 DNA.  
ORGANISM Nostoc sp. PCC 7120  
Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.

REFERENCE 1  
AUTHORS Kaneko,T., Nakamura,Y., Wolk,C.P., Kuritz,T., Sasamoto,S.,  
Watanabe,A., Iriguchi,M., Ishikawa,A., Kawashima,K., Kimura,T.,  
Kishida,Y., Kohara,M., Matsumoto,M., Matsuno,A., Muraki,A.,  
Nakazaki,N., Shimo,S., Sugimoto,M., Takazawa,M., Yamada,M.,  
Yasuda,M. and Tabata,S.  
TITLE Complete genomic sequence of the filamentous nitrogen-fixing  
cyanobacterium Anabaena sp. strain PCC 7120  
JOURNAL DNA Res. 8 (5), 205-213 (2001)  
MEDLINE 21595285  
REFERENCE 2 (bases 1 to 343550)  
AUTHORS Kaneko,T.  
TITLE Direct Submission  
JOURNAL Submitted (02-MAY-2001) Takakazu Kaneko, Kazusa DNA Research  
Institute, The First Laboratory for Plant Gene Research; Yana  
1532-3, Kisarazu, Chiba 292-0812, Japan  
(E-mail:kaneko@kazusa.or.jp,  
URL:http://www.kazusa.or.jp/cyanobase/,  
Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)

FEATURES  
source  
1. .343550  
/organism="Nostoc sp. PCC 7120"  
/db\_xref="taxon:103690"  
/note="synonym:Anabaena sp. PCC7120"  
complement(94. .411)  
/gene="all1719"  
complement(94. .411)  
/gene="all1719"  
/note="ORF ID:all1719  
hypothetical protein"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB73418.1"  
/db\_xref="GI:17130809"  
/translation="MRKKSAILEAVHETAEDLHKAGLMNQTTLREFEHLCLPPIEPL  
EPLQIKEIRSSQVSQAVFARILNISTSTVQKWEIGQKRPASLKLHLVKNRGLNS  
VLY"  
complement(504. .782)  
/gene="as11720"  
complement(504. .782)  
/gene="as11720"  
/note="ORF ID:as11720  
hypothetical protein"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB73419.1"  
/db\_xref="GI:17130810"  
/translation="MKIYKNRTFDRWARKEGLKNLSLCNAVNMAGLYDADLGGGLF  
KKRIAKPGKSGGFRTLIATNNEDVGFLFLAFQKTNAVTLTHMKKKL"  
935. .1507  
/gene="alr1721"  
935. .1507  
/gene="alr1721"  
/note="ORF ID:alr1721"  
/codon\_start=1  
/transl\_table=11  
/product="transcriptional regulator"  
/protein\_id="BAB73420.1"

---

gene  
CDS

gene  
CDS

gene  
CDS

gene  
CDS

gene  
CDS

/db\_xref="GI:17130811"  
/translation="MRHKDNKKNOAICDAEILITANGFADTSMSKIAKAANVSPATI  
YVYFENKEKLNKIYLFVKQEISAEILKGVNQNLSEKAFKMIWNYYKYAIKNPVRF  
AFTEQFANSPMVDRIKDESLSYFKPMLDLFNRGKEKVFKDIPLEIFTAFTFAPLTG  
LIKEHFGALILDEKMLETYYKIAWDAVTN"  
1569. .2504  
/gene="alr1722"  
1569. .2504  
/gene="alr1722"  
/note="ORF ID:alr1722  
probable oxidoreductase"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB73421.1"  
/db\_xref="GI:17130812"  
/translation="MKNEKWAENIPPQKGLAIVTGSSSGIGYETARVLANKQASVI  
IAVRNLDKGNKALAKILQQNKDADVCLKMELDLANLASVKNFENFRKNYLRDLINN  
AGVMIPPYSKTTDCFELQFGTNHLGHFALTGQLLEFLISTEGSRIVNVSSGAHNMGKI  
DFDDLNWEQRSYAKWAGDSKLANLYFTYELDRKLDKNGIDTLVTASHPGWTATELQ  
RTAGGIVKYLNGIVAQDITMGALPTLRAAIEAGLKGAEYFGPNGFMEMRGYPPIKVESN  
ELSKDQALAKKLWVVEKLTDVKFEFNKKAQSAGK"  
complement(2701. .3462)  
/gene="all1723"  
complement(2701. .3462)  
/gene="all1723"  
/note="ORF ID:all1723  
probable aldolase"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB73422.1"  
/db\_xref="GI:17130813"  
/translation="MSVFSRPLPPVFNSEIEERLHRKQRLAAAFRLFGKGFSEGIAG  
HITARDPEFTDHFVNWNPGLGKYFGHIRVSDLVLRDGDVVKGDALINQAAFAIHSQIH  
EARPDVAAAHASIYGKAWSSGLRLLDPLTDQACAFYEDHALLDDYTGVIETSEGK  
RLAETLGNKALILRNHGILTVGHTVDEAAFWYISLERSCQAQLLAEAGQPKIIGHE  
TARLTHSQVGEPRKGFWSFQPLYDRVAREEPELFD"  
complement(3764. .5167)  
/gene="all1724"  
complement(3764. .5167)  
/gene="all1724"  
/note="ORF ID:all1724  
unknown protein"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB73423.1"  
/db\_xref="GI:17130814"  
/translation="MSSTSQGVPLPDLTHPTELIQGSQKITASLLLVFLVIALGIA  
IALMSFALRHNSQOTIFIGEWAIRAQILRGQLHALVLILLVGFFLSSSTLSNRYH  
HWEAKIAQVQSVSGDKLEQIAPQVRVYTOEPTYNTQVNGKIVKNDTQOVSRLLS  
LAGSQIQVKLEQSTDVQGRSAVRADYTDYKVNRLKDINNFFFEIQPPNGYSLLSN  
YKVERDRTRLEPNPCNYSFAFRLEPGQETSERVYQAQGPRWVYNSAGQLLSNFRLL  
TTIANFPGADFASGIVPKERKADGRSTQYTWIFDQNVSVKNPFGVFTNTNPIRQTGII  
PRLLLPALFLWIIILYLSLPISLTNVAIASIYFACLLTLTLYLSRFMDAQLAWTM  
ISLVLLTLTWGLGSTRSASLAAICTIAGAVVPVIGLLVPYSGLTLSVAGLLSVIWL  
VRHWYNWNIWQPEDSKS"  
complement(5492. .5902)  
/gene="all1725"  
complement(5492. .5902)  
/gene="all1725"  
/note="ORF ID:all1725  
unknown protein"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB73424.1"  
/db\_xref="GI:17130815"  
/translation="MFLKPFSGLMVIAASLTNLNGIAITYSGRSAIAEAKSNFYLQVGA  
DQRGFPIALDLASINGTNFTLVQOHGEGIAKRRLHAACSQGRFLTCKKFSLYAANGKIT  
SEDKTEQEIFPKPTPDAASMEIVCRVANNHOKK"  
6321. .6842  
/gene="alr1726"  
6321. .6842  
/gene="alr1726"  
/note="ORF ID:alr1726"

/codon\_start=1
/transl\_table=11
/product="transposase"
/protein\_id="BAB73425.1"
/db\_xref="GI:17130816"
/translation="MKSQHLOATTLAGESVQOYLKEETALNAIAELQDFIDMCPDARE
VRKALAVKLVYQGYLYDEIQKILDSVRSGITGWKQAYEETGIDGLRLNYKGRKSYLNA
KOLQEVLSWLQTKDCWELGELEYQLAFYDVVYESKQSYVDLFEAGISWKKTKVNP
KADENAVAEKKRD"
6919..7431
/gene="alr1727"
6919..7431
/gene="alr1727"
/note="ORF\_ID:alr1727"
/codon\_start=1
/transl\_table=11
/product="transposase"
/protein\_id="BAB73426.1"
/db\_xref="GI:17130817"
/translation="MWGDLGYVWGKSDQEIAPVNVNERDKQTYYGAVDYLLGELVLK
AYDAGNSKNTINYLEYLLANSPDQRLLI FWDGASYHRSKEIRGLDSVNQSLPTEQWK
IHCVRFPAPNCPVQPIEDIWLQAKTWVRRFCALIPSPSHLKWIFEFIRHTTDFDFTL
QMYGYSKIK"
8064..9401
/gene="alr1728"
8064..9401
/gene="alr1728"
/note="ORF\_ID:alr1728"
/hypothetical protein"
/codon\_start=1
/transl\_table=11
/protein\_id="BAB73427.1"
/db\_xref="GI:17130818"
/translation="MTGKNARHNAFLRLVSGNGAAGFSGESRYSLSKEVVIGRDPSCQ
VVLDDAMMYRMVSRRAVVRPVASSVDSKFSWVLCNLNSANGTYLNGORLYGCQELHAG
DRISLGADGQPFLEFYAVAFQPTIITNQVAPLPSAKHPKPDSPVFTQLFPIISTGKD
LTRKAYLVPGLITVVFVLMFATVGRPOANQVIVATIALAAYFYIYRLCGKPKPWV
LVGAALSTLILSLPLNLFIFREVLPGSVPPIDQVSSLTFLVGYFFGAGLMEE
LLKALPILGVLYIALGLPPAWRERVGWEPDLGILLGTASAVGFTLLETGLGVYVPAAS
LOAGTEVGLQVLIARILGLPAGHMAYSGLGYFIFGLAALKPRHAKQILAVGYLSAAAL
HALWNTAGHSNNLLVVGVLSYAFMLMAAILKARALSPTRSQNFATRFLDPKS"
complement(9666..10310)
/gene="all1729"
complement(9666..10310)
/gene="all1729"
/note="ORF\_ID:all1729"
/unknown protein"
/codon\_start=1
/transl\_table=11
/protein\_id="BAB73428.1"
/db\_xref="GI:17130819"
/translation="WTHKTNQSDACHSVVSQSMELDLAALLEPEDAAYPNWPSDDESE
TYFEELERKFAQDVSDDEFTTQAVFYDKLDHLWSGVTPSSYYNHTLSVVVDHLQEN
LHTAFATGIPQNLWLNIAIAATKAAEIFASQQSLGEQLVECVSVLPTWEADDLLIFARPY
AYSMRSEGNLTSTVDNVNQNQWSALSSEVEQAKLSVAIAYAIRQLNDSPSES"
complement(10795..12372)
/gene="all1730"
complement(10795..12372)
/gene="all1730"
/note="ORF\_ID:all1730"
similar to Hetf protein"

Query Match 2.2%; Score 21; DB 1; Length 343550;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 343 CCCAATGTTGCCAAATATGCT 363
| | | | | | | | | | | | | | | | | | | | |
Db 36059 CCCAATGTTGCCAAATATGCT 36079

RESULT 11
HUMLDHX3

LOCUS HUMLDHX3 188 bp DNA linear PRI 08-MAY-2000
DEFINITION Homo sapiens lactate dehydrogenase-C (LDHC) gene, exon 4.
ACCESSION M24512
VERSION M24512.1 GI:187069
KEYWORDS
SEGMENT 3 of 6
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 188)
AUTHORS Takano,T. and Li,S.S.
TITLE Human testicular lactate dehydrogenase-C gene is interrupted by six
introns at positions homologous to those of LDH-A (muscle) and
LDH-B (heart) genes
JOURNAL Biochem. Biophys. Res. Commun. 159 (2), 579-583 (1989)
MEDLINE 89193640
PUBMED 2930531
COMMENT Draft entry and computer-readable sequence for [1] kindly submitted
by S.S.-L.Li, 28-APR-1989.
FEATURES
source
Location/Qualifiers
1..188
/organism="Homo sapiens"
/db\_xref="taxon:9606"
/map="11p15.5-p14.3"
/clone="hC[201,203,210,211,303]"
/tissue\_type="testis"
<1..4
/gene="LDHC"
/note="C"
5..178
/gene="LDHC"
/note="G00-120-150"
/number=4
179..>188
/gene="LDHC"
/note="D"
BASE COUNT 50 a 29 c 48 g 61 t
ORIGIN
Query Match 2.1%; Score 20; DB 9; Length 188;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 547 ATGGAGAACATGGTGATTC 566
| | | | | | | | | | | | | | | | | | | | |
Db 154 ATGGAGAACATGGTGATTC 173

RESULT 12
AF401096
LOCUS AF401096 811 bp mRNA linear PRI 02-AUG-2002
DEFINITION Homo sapiens lactate dehydrogenase C (LDHC) mRNA, complete
sequence, alternatively spliced.
ACCESSION AF401096
VERSION AF401096.1 GI:22086508
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 811)
AUTHORS Koslowski,M., Tuereci,O. and Sahin,U.
TITLE Identification of new splice variants of testis-specific lactate
dehydrogenase C expressed exclusively in tumors by data-mining of
GenBank
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 811)
AUTHORS Koslowski,M., Tuereci,O. and Sahin,U.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2001) Department of Internal Medicine
III/Hematology, University Mainz, Obere Zahlbacherstr. 63, Mainz
55131, Germany



FEATURES  
source Location/Qualifiers  
1. .811  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11p15.5-p15.3"  
/tissue\_type="lung cancer"  
1. .811  
/gene="LDHC"  
/note="lactate dehydrogenase C variant 5; tumor specific;  
coding region not determined; lacks exons 2, 6 and 7"  
BASE COUNT 235 a 135 c 184 g 257 t  
ORIGIN  
Query Match 2.1%; Score 20; DB 9; Length 811;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 547 ATTGGAGAACATGGTGATTC 566  
Db 508 ATTGGAGAACATGGTGATTC 527  
RESULT 13  
AF401095  
LOCUS AF401095 879 bp mRNA linear PRI 02-AUG-2002  
DEFINITION Homo sapiens lactate dehydrogenase C (LDHC) mRNA, complete  
sequence, alternatively spliced.  
ACCESSION AF401095  
VERSION AF401095  
KEYWORDS AF401095.1 GI:22086507  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 879)  
AUTHORS Koslowski,M., Tuereci,O. and Sahin,U.  
TITLE Identification of new splice variants of testis-specific lactate  
dehydrogenase C expressed exclusively in tumors by data-mining of  
GenBank  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 879)  
AUTHORS Koslowski,M., Tuereci,O. and Sahin,U.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUL-2001) Department of Internal Medicine  
III/Hematology, University Mainz, Obere Zahlbacherstr. 63, Mainz  
55131, Germany  
FEATURES  
source Location/Qualifiers  
1. .879  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11p15.5-p15.3"  
/tissue\_type="lung cancer"  
1. .879  
/gene="LDHC"  
/note="lactate dehydrogenase C variant 4; testis specific;  
coding region not determined; lacks exons 3 and 4"  
BASE COUNT 254 a 139 c 210 g 276 t  
ORIGIN  
Query Match 2.1%; Score 20; DB 9; Length 879;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 547 ATTGGAGAACATGGTGATTC 566  
Db 334 ATTGGAGAACATGGTGATTC 353  
RESULT 14  
AF401097  
LOCUS AF401097 1047 bp mRNA linear PRI 02-AUG-2002

DEFINITION Homo sapiens lactate dehydrogenase C (LDHC) mRNA, complete  
sequence, alternatively spliced.  
ACCESSION AF401097  
VERSION AF401097.1 GI:22086509  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1047)  
AUTHORS Koslowski,M., Tuereci,O. and Sahin,U.  
TITLE Identification of new splice variants of testis-specific lactate  
dehydrogenase C expressed exclusively in tumors by data-mining of  
GenBank  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1047)  
AUTHORS Koslowski,M., Tuereci,O. and Sahin,U.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUL-2001) Department of Internal Medicine  
III/Hematology, University Mainz, Obere Zahlbacherstr. 63, Mainz  
55131, Germany  
FEATURES  
source Location/Qualifiers  
1. .1047  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11p15.5-p15.3"  
/tissue\_type="lung cancer"  
1. .1047  
/gene="LDHC"  
/note="lactate dehydrogenase C variant 6; tumor specific;  
coding region not determined; lacks exon 7"  
BASE COUNT 305 a 172 c 243 g 327 t  
ORIGIN  
Query Match 2.1%; Score 20; DB 9; Length 1047;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 547 ATTGGAGAACATGGTGATTC 566  
Db 626 ATTGGAGAACATGGTGATTC 645  
RESULT 15  
AF401094  
LOCUS AF401094 1053 bp mRNA linear PRI 02-AUG-2002  
DEFINITION Homo sapiens lactate dehydrogenase C (LDHC) mRNA, complete  
sequence, alternatively spliced.  
ACCESSION AF401094  
VERSION AF401094.1 GI:22086506  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1053)  
AUTHORS Koslowski,M., Tuereci,O. and Sahin,U.  
TITLE Identification of new splice variants of testis-specific lactate  
dehydrogenase C expressed exclusively in tumors by data-mining of  
GenBank  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1053)  
AUTHORS Koslowski,M., Tuereci,O. and Sahin,U.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUL-2001) Department of Internal Medicine  
III/Hematology, University Mainz, Obere Zahlbacherstr. 63, Mainz  
55131, Germany  
FEATURES  
source Location/Qualifiers  
1. .1053  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"

```
gene /map="11p15.5-p15.3"
      /tissue type="lung cancer"
      1..1053
      /gene="LDHC"
      /note="lactate dehydrogenase C variant 3; tumor specific;
      coding region not determined; lacks exon 3"
BASE COUNT 305 a 176 c 246 g 326 t
ORIGIN
Query Match 2.1%; Score 20; DB 9; Length 1053;
Best Local Similarity 100.0%; Pred.No. 27;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 547 ATTGGAGAACATGGTGATTC 566
   |||||
Db 508 ATTGGAGAACATGGTGATTC 527
```

Search completed: July 31, 2003, 12:44:47  
Job time : 2881 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 10:16:32 ; Search time 278 Seconds  
(without alignments)  
7873.893 Million cell updates/sec

Title: US-09-992-430B-21  
Perfect score: 972  
Sequence: 1 atgtccaagatacaaaagtc.....ttgagcatctggagatcaat 972

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N\_Geneseq 101002:\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
1	20	2.1 1234 9	Sequence encoding
2	19	2.0 1680 24	Human cDNA encodin
3	19	2.0 1755 22	Human secreted pro
4	19	2.0 1759 22	Human secreted pro
5	19	2.0 4068 22	Mouse p97 (mp97) p
6	19	2.0 4158 22	DNA encoding membr
7	19	2.0 5313 23	Drosophila melanog
8	19	2.0 33140 22	Genomic fragment #
9	19	2.0 1082138 21	Arabidopsis thalia



C	83	17	1.7	1884	22	AAH64742	Human secreted pro
	84	17	1.7	1954	22	AAD20950	Wheat arginyl-tRNA
	85	17	1.7	1992	24	ABK63809	Rat sequence diffe
	86	17	1.7	2387	14	AAQ48438	Brh-I cDNA, Braco
	87	17	1.7	2387	17	AAT45675	Wasp Brh-1 toxin g
	88	17	1.7	2387	22	AAS05984	Parasitic wasp ins
	89	17	1.7	2642	23	ABL10304	Drosophila melanog
	90	17	1.7	3019	22	AAH54991	S. epidermidis gen
	91	17	1.7	3172	21	AAA95794	Apoptosis related
	92	17	1.7	3532	22	AAH15730	Human cDNA sequenc
	93	17	1.7	3573	23	ABL05120	Drosophila melanog
	94	17	1.7	3799	24	ABK70290	Human lung cancer
	95	17	1.7	3799	24	ABK24337	DNA encoding human
	96	17	1.7	4082	24	ABK09788	Human ovarian tumo
	97	17	1.7	4093	22	AAH54429	S. epidermidis gen
	98	17	1.7	4239	23	AAS64518	DNA encoding novel
	99	17	1.7	4715	22	AAK84674	Human immune/haema
	100	17	1.7	4881	19	AAV52258	Streptococcus pneu
	101	17	1.7	4902	23	AAS85161	DNA encoding novel
	102	17	1.7	5107	24	ABL33066	Human immune syste
	103	17	1.7	5279	23	AAS85168	DNA encoding novel
	104	17	1.7	5326	23	ABL29084	Drosophila melanog
	105	17	1.7	5600	13	AAQ21678	pHCV-31 - recombin
	106	17	1.7	5600	13	AAQ22963	Sequence of pHCV-3
	107	17	1.7	5600	14	AAQ38266	HCV CKS-33-BCD fus
	108	17	1.7	5600	14	AAQ38251	HCV CKS-33-BCD fus
	109	17	1.7	5600	14	AAQ38236	HCV CKS-33-BCD exp
	110	17	1.7	5600	15	AAQ62663	pHCV-31 plasmid ex
	111	17	1.7	5600	22	AAF32219	HCV recombinant an
	112	17	1.7	6046	23	ABL07762	Drosophila melanog
	113	17	1.7	6560	15	AAQ54629	Genomic region con
	114	17	1.7	6765	22	AAH90108	Human bone marrow
	115	17	1.7	6798	22	AAI99051	Human excretory re
	116	17	1.7	6798	22	AAI63401	Human kidney relat
	117	17	1.7	6798	23	ABK44028	Genomic DNA encodi
	118	17	1.7	7491	24	ABL33585	Human immune syste
	119	17	1.7	7536	23	ABL08520	Drosophila melanog
	120	17	1.7	8022	18	AAT61982	Yeast rad3 gene.
	121	17	1.7	9001	22	AAK84675	Human immune/haema
	122	17	1.7	9179	23	ABL03688	Drosophila melanog
	123	17	1.7	9771	22	AAK71552	Human immune/haema
	124	17	1.7	9979	22	AAK81194	Human immune/haema
	125	17	1.7	10382	22	AAK67484	Human immune/haema
	126	17	1.7	12177	23	ABL03724	Drosophila melanog
	127	17	1.7	13673	22	AAK76548	Human immune/haema
	128	17	1.7	13673	22	AAK81195	Human immune/haema
	129	17	1.7	14676	22	AAK71553	Human immune/haema
	130	17	1.7	14962	22	AAS29980	Human lung antigen
	131	17	1.7	15496	22	AAK71549	Human immune/haema
	132	17	1.7	16766	24	ABL34157	Human immune syste
	133	17	1.7	16983	22	AAK85126	Human immune/haema
	134	17	1.7	19191	22	AAK67485	Human immune/haema
	135	17	1.7	21724	22	AAS26629	Human immune/haema
	136	17	1.7	21724	22	AAK86125	Human genomic DNA
	137	17	1.7	21727	22	AAS26630	Human immune/haema
	138	17	1.7	21727	22	AAK86126	Human immune/haema
	139	17	1.7	23054	22	AAK84676	Human immune/haema
	140	17	1.7	32217	22	AAS41742	Genomic sequence #
	141	17	1.7	34488	22	AAF97854	Human neuroblastom
	142	17	1.7	50925	21	AAA81487	N. meningitidis pa
	143	17	1.7	59560	22	AAK71550	Human immune/haema
	144	17	1.7	59560	22	AAK80620	Human immune/haema
	145	17	1.7	114793	22	AAD08215	Human genome from
	146	17	1.7	121162	21	AAC66548	Human kinesin-like
	147	17	1.7	128600	24	ABK83461	Human cDNA differe
	148	17	1.7	309400	22	AAH68534	C glutamic codin
	149	17	1.7	349980	21	AAF21610	Neisseria meningit
	150	17	1.7	534720	19	AAV30458	Rhizobium species
	151	17	1.7	536165	19	AAV30459	Rhizobium species
	152	17	1.7	1230025	20	AAX91990	Nucleotide sequenc
	153	17	1.7	1437668	21	AAA81490	N. meningitidis B
	154	16	1.6	81	18	AAT57932	L-selectin orphan
	155	16	1.6	87	16	AAT19428	Human gene signatu

C	156	16	1.6	176	23	ABV38418	Human prostate exp
	157	16	1.6	233	21	AAC21576	Human secreted pro
	158	16	1.6	240	22	AAK23943	Human brain expres
	159	16	1.6	240	24	ABN66494	Streptococcus poly
	160	16	1.6	251	19	AAK11049	Human biallelic po
	161	16	1.6	251	19	AAK11050	Human biallelic po
	162	16	1.6	261	14	AAQ39685	Expressed Sequence
	163	16	1.6	289	21	AAZ49263	Human hydrolase ho
	164	16	1.6	300	13	AAQ25393	Antisense DNA comp
	165	16	1.6	300	13	AAQ25394	RI alpha 100 N-ter
	166	16	1.6	300	17	AAK34984	A fragment of the
	167	16	1.6	300	17	AAK34989	A fragment of the
	168	16	1.6	300	18	AAK77365	RI-alpha antisense
	169	16	1.6	300	18	AAK77366	RI-alpha 5'-termin
	170	16	1.6	300	21	AAZ59438	CAMP dependent kin
	171	16	1.6	300	21	AAZ59439	CAMP dependent kin
	172	16	1.6	310	22	AAK76777	Human immune/haema
	173	16	1.6	316	20	AAZ21544	Fragment of progre
	174	16	1.6	316	20	AAZ25859	Clone PEGen-28 seq
	175	16	1.6	321	21	AAK39327	S. lactificifex 16S
	176	16	1.6	324	24	AAK37377	DLP10-1 adenylate
	177	16	1.6	326	24	ABQ55704	Human ovarian anti
	178	16	1.6	328	23	ABV08516	Human prostate exp
	179	16	1.6	340	22	AAI89619	Human polynucleoti
	180	16	1.6	345	24	ABL62008	Colon adenocarcino
	181	16	1.6	345	24	ABL62522	Colon adenocarcino
	182	16	1.6	354	21	AAC56838	Pinus radiata tran
	183	16	1.6	369	22	AAI85767	Human polynucleoti
	184	16	1.6	426	22	AAH34655	Human colon cancer
	185	16	1.6	435	22	AAS30998	Human diagnostic a
	186	16	1.6	438	21	AAC93763	Cat flea hindgut a
	187	16	1.6	438	24	ABQ90494	M. capsulatus gene
	188	16	1.6	448	21	AAC56450	Pinus radiata tran
	189	16	1.6	450	21	AAF16504	Human prostate can
	190	16	1.6	452	20	AAV88641	EST clone GY520.
	191	16	1.6	462	22	ABA46157	Human breast cell
	192	16	1.6	462	22	AAK30360	Human bone marrow
	193	16	1.6	462	22	AAI04738	Probe #4729 used t
	194	16	1.6	462	24	ABQ04951	Human genome-deriv
	195	16	1.6	464	21	AAA42393	Human secreted exp
	196	16	1.6	465	22	AAH98232	Human EST-derived
	197	16	1.6	477	23	ABV61300	Human prostate exp
	198	16	1.6	477	23	AAS5093	DNA encoding novel
	199	16	1.6	484	22	AAF93377	Spinal cord tissue
	200	16	1.6	495	22	AAS40318	DNA encoding human
	201	16	1.6	495	22	AAI03918	Human reproductive
	202	16	1.6	504	23	ABV61372	Human prostate exp
	203	16	1.6	508	21	AAK07611	Human secreted pro
	204	16	1.6	514	22	AAH06737	Human cDNA clone (
	205	16	1.6	524	24	ABQ51908	Oligonucleotide fo
	206	16	1.6	524	24	ABQ51909	Oligonucleotide fo
	207	16	1.6	528	15	AAQ73887	Borrelia 297 antig
	208	16	1.6	535	23	ABV51010	Human prostate exp
	209	16	1.6	553	24	ABQ34610	Oligonucleotide fo
	210	16	1.6	553	24	ABQ34611	Oligonucleotide fo
	211	16	1.6	554	22	AAK11282	Human brain expres
	212	16	1.6	563	24	ABN62146	Human cancer relat
	213	16	1.6	573	22	AAF29011	B burgdorferi ospC
	214	16	1.6	579	15	AAQ73862	Borrelia 297 antig
	215	16	1.6	579	21	AAA39326	S. lactificifex 16S
	216	16	1.6	580	22	AAI02557	Human reproductive
	217	16	1.6	595	22	AAK61803	Human immune/haema
	218	16	1.6	595	22	AAS34776	CDNA encoding nove
	219	16	1.6	615	21	AAZ49233	Human hydrolase ho
	220	16	1.6	621	22	AAS45236	CDNA encoding nove
	221	16	1.6	623	22	AAS45048	CDNA encoding nove
	222	16	1.6	645	22	AAS40380	DNA encoding human
	223	16	1.6	645	22	AAS40383	Human reproductive
	224	16	1.6	645	22	AAI03983	Human reproductive
	225	16	1.6	645	22	AAI03986	Human reproductive
	226	16	1.6	663	11	AAQ05615	Insert from pARC 1
	227	16	1.6	694	21	AAZ53850	Neisseria gonorrhoe
	228	16	1.6	700	22	AAH92618	Human inflammatory

C 229	16	1.6	708	24	ABQ37310	Oligonucleotide fo	C 302	16	1.6	1538	20	AAZ21530	Progression elevat
C 230	16	1.6	708	24	ABQ37311	Oligonucleotide fo	C 303	16	1.6	1544	18	AAT86839	CDNA encoding rat
C 231	16	1.6	738	22	AAH05900	Human cDNA clone (	304	16	1.6	1573	24	ABN86082	Human caudal actio
C 232	16	1.6	744	21	AAZ53851	Neisseria meningit	305	16	1.6	1588	24	ABQ54276	Human ovarian anti
C 233	16	1.6	749	22	AAI95692	Human neuroblastom	C 306	16	1.6	1614	22	AAH98192	Human EST-derived
C 234	16	1.6	761	22	AAI96716	Human neuroblastom	C 307	16	1.6	1632	23	ABV24429	Human prostate exp
C 235	16	1.6	762	24	ABQ68799	Listeria monocytog	308	16	1.6	1668	22	AAH43384	CDNA sequence enco
C 236	16	1.6	812	21	AAZ35438	Arabidopsis thalia	C 309	16	1.6	1683	23	AAS54012	Klebsiella pneumon
C 237	16	1.6	812	24	ABQ35900	Oligonucleotide fo	C 310	16	1.6	1688	22	AAK68785	Human immune/haema
C 238	16	1.6	812	24	ABQ35901	Oligonucleotide fo	311	16	1.6	1709	13	AAQ29277	Encodes secreted f
C 239	16	1.6	815	21	AAZ39440	Arabidopsis thalia	312	16	1.6	1714	21	AAC93536	Cat flea HMT Na/K
C 240	16	1.6	825	22	AAH65959	C glutamicum codin	C 313	16	1.6	1714	21	AAC93537	Cat flea HMT Na/K
C 241	16	1.6	829	20	AAZ13623	Enterococcus faeca	C 314	16	1.6	1737	19	AAV29053	Alcohol and/or ald
C 242	16	1.6	873	23	AAS87910	DNA encoding novel	315	16	1.6	1741	22	AAK56769	Human immune/haema
C 243	16	1.6	879	22	AAF27597	Mevalonate pathway	316	16	1.6	1750	22	AAH14806	Human cDNA sequenc
C 244	16	1.6	891	24	ABN69039	Streptococcus poly	C 317	16	1.6	1751	24	ABL90377	Human polynucleoti
C 245	16	1.6	895	23	ABV25466	Human prostate exp	318	16	1.6	1785	19	AAV41902	Nucleotide sequenc
C 246	16	1.6	941	24	ABK12491	Arabidopsis thalia	319	16	1.6	1788	23	AAS51607	Staphylococcus aur
C 247	16	1.6	960	19	AAV53346	DNA encoding a Sta	320	16	1.6	1806	19	AAV41901	Nucleotide sequenc
C 248	16	1.6	972	23	ABL22115	Drosophila melanog	321	16	1.6	1815	23	AAS54790	Staphylococcus aur
C 249	16	1.6	978	21	AAC93338	Cat flea HMT Na/K	322	16	1.6	1828	23	ABL20439	Drosophila melanog
C 250	16	1.6	978	21	AAC93339	Cat flea HMT Na/K	323	16	1.6	1854	22	AAI59167	Human polynucleoti
C 251	16	1.6	987	23	AAS56137	Salmonella typhi D	324	16	1.6	1948	21	AAC81726	Human secreted pro
C 252	16	1.6	1001	21	AAC57469	Arachidonic acid m	325	16	1.6	1949	22	AAH75170	Nucleotide sequenc
C 253	16	1.6	1023	23	ABL27951	Drosophila melanog	326	16	1.6	1967	22	ABA44267	Human breast cell
C 254	16	1.6	1036	21	AAF16278	Human prostate can	327	16	1.6	1967	22	ABA54718	Human foetal liver
C 255	16	1.6	1066	21	AAA26671	Candida albicans p	328	16	1.6	1967	22	ABA24498	Probe #2964 for ge
C 256	16	1.6	1066	22	AAF76852	Human secreted pro	329	16	1.6	1967	22	AAK03000	Human brain expres
C 257	16	1.6	1068	24	ABK73448	Bacillus lichenifo	330	16	1.6	1967	22	AAK28450	Human bone marrow
C 258	16	1.6	1084	22	AAI61153	Human polynucleoti	331	16	1.6	1967	22	AAI13018	Probe #2951 for ge
C 259	16	1.6	1089	24	ABN70774	Streptococcus poly	332	16	1.6	1967	22	AAI34371	Probe #3057 used t
C 260	16	1.6	1107	22	AAF29022	Borrelia sp chimex	333	16	1.6	1967	22	AAI02933	Probe #2924 used t
C 261	16	1.6	1110	22	AAF27709	Human transport pr	334	16	1.6	1967	22	AAS00836	Human cDNA clone H
C 262	16	1.6	1113	22	AAF29019	Borrelia sp chimex	335	16	1.6	1967	24	ABS02957	Human genome-deriv
C 263	16	1.6	1121	21	AAZ89015	A. ambisexualis Aa	336	16	1.6	1982	23	ABL27901	Drosophila melanog
C 264	16	1.6	1125	24	ABN66495	Streptococcus poly	C 337	16	1.6	1997	24	ABI99512	Mouse ischaemic co
C 265	16	1.6	1128	22	AAF29014	Borrelia sp chimex	338	16	1.6	2025	22	AAI92899	Human polynucleoti
C 266	16	1.6	1137	22	AAF29016	Borrelia sp chimex	C 339	16	1.6	2054	13	AAQ20462	Aspergillus tubige
C 267	16	1.6	1176	24	ABL39743	Human NS cDNA sequ	C 340	16	1.6	2054	20	AAK00973	Aspergillus tubige
C 268	16	1.6	1179	22	AAF29038	Borrelia sp chimex	341	16	1.6	2095	8	AAW70692	Sequence of human
C 269	16	1.6	1185	22	AAF29034	Borrelia sp chimex	342	16	1.6	2095	17	AAT27029	Atrial natriuretic
C 270	16	1.6	1196	21	AAC53987	Arabidopsis thalia	343	16	1.6	2096	21	AAC50354	Arabidopsis thalia
C 271	16	1.6	1197	22	AAF29032	Borrelia sp chimex	C 344	16	1.6	2102	22	AAF33214	Human secreted pro
C 272	16	1.6	1209	22	AAF29042	Borrelia sp chimex	C 345	16	1.6	2129	16	AAQ86899	C. albicans FKSl h
C 273	16	1.6	1239	23	AAS54243	Pseudomonas aerugi	346	16	1.6	2147	22	AAK91479	Human digestive sy
C 274	16	1.6	1267	23	ABL21425	Drosophila melanog	347	16	1.6	2147	22	AAI57789	Human colorectal c
C 275	16	1.6	1277	21	AAC40265	Arabidopsis thalia	348	16	1.6	2187	23	AAS65344	DNA encoding novel
C 276	16	1.6	1277	21	AAC48601	Arabidopsis thalia	349	16	1.6	2187	23	AAS71960	DNA encoding novel
C 277	16	1.6	1292	22	AAF90003	Nucleotide sequenc	350	16	1.6	2187	23	AAS72665	DNA encoding novel
C 278	16	1.6	1299	24	ABK73383	Bacillus lichenifo	C 351	16	1.6	2187	23	AAS82384	DNA encoding novel
C 279	16	1.6	1316	24	ABQ65390	Human cypial gene	352	16	1.6	2187	23	AAS90134	DNA encoding novel
C 280	16	1.6	1331	13	AAQ24681	H23-ETA-S antigen	353	16	1.6	2236	22	ABA07878	Human ovarian and
C 281	16	1.6	1345	22	ABA49421	Human breast cell	354	16	1.6	2236	22	AAI03695	Human reproductive
C 282	16	1.6	1345	22	ABA67334	Human foetal liver	C 355	16	1.6	2236	22	AAK66847	Human immune/haema
C 283	16	1.6	1345	22	ABA34425	Probe #12891 for g	C 356	16	1.6	2236	22	AAK69056	Human immune/haema
C 284	16	1.6	1345	22	AAK15766	Human brain expres	357	16	1.6	2236	22	AAK72813	Human immune/haema
C 285	16	1.6	1345	22	AAK41508	Human bone marrow	C 358	16	1.6	2260	21	AAZ48244	Human oxidoreducta
C 286	16	1.6	1345	22	AAI22256	Probe #12189 for g	C 359	16	1.6	2260	21	AAZ49224	Human hydrolase ho
C 287	16	1.6	1345	22	AAI47547	Probe #16233 used	C 360	16	1.6	2266	22	AAS25838	Human cDNA encodin
C 288	16	1.6	1345	22	AAI07952	Probe #7943 used t	361	16	1.6	2268	23	AAS52421	E. coli DNA for ce
C 289	16	1.6	1345	24	ABS15512	Human genome-deriv	C 362	16	1.6	2292	24	ABQ70772	Listeria monocytog
C 290	16	1.6	1350	23	AAS89168	DNA encoding novel	363	16	1.6	2310	24	AAS94752	Human DNA sequence
C 291	16	1.6	1412	21	AAC76222	Human ORFX ORF177	364	16	1.6	2394	22	AAC85095	Atherosclerosis-as
C 292	16	1.6	1475	9	AAN80638	Sequence of cDNA e	C 365	16	1.6	2417	22	AAH14267	Human cDNA sequenc
C 293	16	1.6	1483	24	ABL90081	Human polynucleoti	366	16	1.6	2471	24	AAI72320	ISIGP-3 cDNA. Hom
C 294	16	1.6	1484	22	AAS27394	CDNA encoding nove	C 367	16	1.6	2490	23	AAS86902	DNA encoding novel
C 295	16	1.6	1484	23	ABK43598	DNA encoding novel	C 368	16	1.6	2513	22	AAK64965	Human immune/haema
C 296	16	1.6	1491	24	ABN91403	Staphylococcus epi	C 369	16	1.6	2520	14	AAQ40580	Glutaryl-7-amino c
C 297	16	1.6	1521	24	ABN81317	Arabidopsis thalia	C 370	16	1.6	2550	23	ABL20341	Drosophila melanog
C 298	16	1.6	1524	23	AAS75643	DNA encoding novel	371	16	1.6	2580	22	AAH54959	S. epidermidis gen
C 299	16	1.6	1524	23	AAS83665	DNA encoding novel	C 372	16	1.6	2591	22	AAH18566	Human cDNA sequenc
C 300	16	1.6	1534	21	AAC33954	Arabidopsis thalia	C 373	16	1.6	2619	21	AAC77768	Human cancer assoc
C 301	16	1.6	1536	24	ABK84287	Human cDNA differe	374	16	1.6	2655	23	AAS71962	DNA encoding novel

375	16	1.6	2655	23	AAS72667	DNA encoding novel	448	16	1.6	6746	23	ABL14122	Drosophila melanog
376	16	1.6	2655	23	AAS77516	DNA encoding novel	449	16	1.6	6891	23	AAS85215	DNA encoding novel
377	16	1.6	2655	23	AAS80016	DNA encoding novel	450	16	1.6	6920	22	AAH98318	Chicken EST-derive
378	16	1.6	2668	23	ABL09903	Drosophila melanog	451	16	1.6	7032	23	AAS85436	DNA encoding novel
379	16	1.6	2685	24	ABK83742	Human cDNA differe	c 452	16	1.6	7033	22	AAS46522	Tumour suppressor
380	16	1.6	2702	22	AAS60836	Human cancer agent	453	16	1.6	8178	24	ABL32087	Human immune syste
381	16	1.6	2702	22	AAS60840	Human cancer agent	454	16	1.6	8181	24	ABL41001	Human neuregulin 2
382	16	1.6	2795	23	ABL16644	Drosophila melanog	455	16	1.6	8342	24	ABL32500	Human myosin IXa c
383	16	1.6	2853	22	AAF33096	Human secreted pro	456	16	1.6	8473	20	AAH25487	Human immune syste
384	16	1.6	2856	22	AAH54076	S. epidermidis gen	c 457	16	1.6	8483	23	ABL25370	Drosophila melanog
385	16	1.6	2926	23	ABL14123	Drosophila melanog	458	16	1.6	9164	20	AAH82259	Beta-domain delete
386	16	1.6	2982	22	AAH54191	S. epidermidis gen	459	16	1.6	9496	22	ABA16328	Human nervous syst
387	16	1.6	2991	19	AAV52588	Vascular endotheli	460	16	1.6	10455	20	AAH02095	Glycine max protop
388	16	1.6	2994	23	AAS85147	DNA encoding novel	461	16	1.6	10598	23	ABL27241	Drosophila melanog
389	16	1.6	3023	23	ABL27950	Drosophila melanog	462	16	1.6	10629	22	AAH28279	Nucleotide sequenc
390	16	1.6	3036	24	ABK84001	Human cDNA differe	463	16	1.6	10856	24	ABL32457	Human immune syste
391	16	1.6	3102	21	AAC75441	Human ORFX ORF996	c 464	16	1.6	10966	22	AAF32380	Hordeum vulgare L.
392	16	1.6	3133	22	AAH14528	Human cDNA sequenc	465	16	1.6	11846	20	AAH82261	Factor VIII protei
393	16	1.6	3227	22	AAH54510	S. epidermidis gen	466	16	1.6	12010	23	ABL06138	Drosophila melanog
394	16	1.6	3367	22	AAK52413	Human polynucleoti	467	16	1.6	12022	20	AAH82260	Factor VIII protei
395	16	1.6	3377	22	AAF72801	Secreted protein g	c 468	16	1.6	13525	23	ABL27900	Drosophila melanog
396	16	1.6	3377	22	AAF72828	Secreted protein g	469	16	1.6	14063	22	AAH82933	Drosophila melanog
397	16	1.6	3382	23	ABL22114	Drosophila melanog	470	16	1.6	14063	22	AAH82934	Human immune/haema
398	16	1.6	3429	22	ABA08830	Human protein kina	471	16	1.6	15231	24	ABK63598	Human immune/haema
399	16	1.6	3465	23	ABL20199	Drosophila melanog	472	16	1.6	15366	22	AAH28623	Rat sequence diffe
400	16	1.6	3503	23	ABL25371	Drosophila melanog	c 473	16	1.6	15667	23	ABL27240	Genomic sequence #
401	16	1.6	3636	23	ABL11340	Drosophila melanog	474	16	1.6	16750	22	AAS46314	Drosophila melanog
402	16	1.6	3825	23	ABV25634	Human prostate exp	475	16	1.6	16750	24	ABL32521	Tumour suppressor
403	16	1.6	3900	23	AAS64300	DNA encoding novel	c 476	16	1.6	16857	23	ABL28687	Human immune syste
404	16	1.6	4002	22	AAH15583	Human cDNA sequenc	c 477	16	1.6	17687	22	AAK64966	Drosophila melanog
405	16	1.6	4024	23	ABL15060	Drosophila melanog	c 478	16	1.6	17687	22	AAK71665	Human immune/haema
406	16	1.6	4154	22	AAD12785	Human nuclear horm	c 479	16	1.6	17689	23	ABL06316	Human immune/haema
407	16	1.6	4207	22	ABA17738	Human nervous syst	c 480	16	1.6	17959	24	ABL54342	Drosophila melanog
408	16	1.6	4208	22	ABA17737	Human nervous syst	481	16	1.6	17959	24	ABL32575	Chemically treated
409	16	1.6	4208	22	ABA17739	Human nervous syst	c 482	16	1.6	17979	22	AAK64964	Human immune syste
410	16	1.6	4230	23	AAS76399	DNA encoding novel	c 483	16	1.6	17979	22	AAK71664	Human immune/haema
411	16	1.6	4231	23	ABL22756	Drosophila melanog	c 484	16	1.6	18439	23	ABL17911	Drosophila melanog
412	16	1.6	4270	22	AAH43377	DNA sequence encod	c 485	16	1.6	18449	22	ABA20913	Human nervous syst
413	16	1.6	4287	24	ABN79843	Fungal ZBC gene se	c 486	16	1.6	18449	22	AAS27797	DNA encoding novel
414	16	1.6	4299	24	ABN79813	Fungal ZBC gene se	c 487	16	1.6	19209	22	AAK70154	Human immune/haema
415	16	1.6	4341	23	ABL20396	Drosophila melanog	c 488	16	1.6	19209	22	AAK86103	Human immune/haema
416	16	1.6	4607	24	ABK46540	DNA encoding human	c 489	16	1.6	20284	21	AAC69139	Human ABC1 gene ex
417	16	1.6	4620	23	AAS85216	DNA encoding novel	490	16	1.6	20829	23	ABL14502	Drosophila melanog
418	16	1.6	4679	22	AAF75865	Histidine protein k	491	16	1.6	21382	23	AAS59557	Propionibacterium
419	16	1.6	4711	21	AAC76508	Human protein enco	c 492	16	1.6	21423	22	AAL36470	Human musculoskele
420	16	1.6	4715	22	AAH99638	Human ORFX ORF2063	493	16	1.6	21729	23	ABL06154	Drosophila melanog
421	16	1.6	4722	23	ABL21424	Drosophila melanog	c 494	16	1.6	22081	22	AAF97862	Human neuroblastom
422	16	1.6	4858	23	ABL20437	Drosophila melanog	495	16	1.6	27423	23	ABL04268	Drosophila melanog
423	16	1.6	4972	23	ABL20340	Drosophila melanog	c 496	16	1.6	31208	24	ABK52899	Genomic DNA encodi
424	16	1.6	4989	23	ABL09902	Drosophila melanog	c 497	16	1.6	32190	22	AAS29937	Human lung antigen
425	16	1.6	5004	23	ABL18692	Drosophila melanog	c 498	16	1.6	32190	22	AAH89689	Human digestive sy
426	16	1.6	5145	19	AAV45185	Nucleotide sequenc	499	16	1.6	32193	22	AAS29936	Human lung antigen
427	16	1.6	5203	22	AAS46473	Tumour suppressor	500	16	1.6	32249	22	AAS29938	Human lung antigen
428	16	1.6	5203	24	ABN80188	Human chemically m	501	16	1.6	34548	24	ABL70603	Chemically treated
429	16	1.6	5311	22	ABA21074	Human nervous syst	502	16	1.6	40766	23	ABL28686	Drosophila melanog
430	16	1.6	5311	22	AAH07486	Human reproductive	503	16	1.6	42429	23	ABL17910	Drosophila melanog
431	16	1.6	5339	24	ABA97074	Human fank cDNA.	504	16	1.6	43058	24	ABN97455	Gene #3953 used to
432	16	1.6	5465	23	ABL20198	Drosophila melanog	505	16	1.6	43058	24	ABL64982	Lung cancer relate
433	16	1.6	5694	22	AAS23416	Candida albicans e	506	16	1.6	43058	24	ABL65219	Lung cancer relate
434	16	1.6	5774	22	AAS46652	Tumour suppressor	c 507	16	1.6	49767	21	AAH81458	N. meningitidis pa
435	16	1.6	5774	24	ABL33163	Human immune syste	c 508	16	1.6	49999	24	ABO90984	M. capsulatus gene
436	16	1.6	5788	24	ABK39942	Human chemically p	c 509	16	1.6	55795	24	ABN95045	Gene #1543 used to
437	16	1.6	5874	23	AAS85420	DNA encoding novel	c 510	16	1.6	55795	24	ABL68242	Kidney cancer rela
438	16	1.6	6160	24	ABL70233	Chemically treated	c 511	16	1.6	55795	24	ABL68484	Kidney cancer rela
439	16	1.6	6160	24	ABK31272	Signal transductio	c 512	16	1.6	55795	24	ABL68863	Kidney cancer rela
440	16	1.6	6228	23	ABL11144	Drosophila melanog	c 513	16	1.6	58407	19	AAV21210	Methanococcus jann
441	16	1.6	6286	24	ABL60158	Human MUC1 encodin	c 514	16	1.6	112190	22	AAH44801	Human GPCR protein
442	16	1.6	6363	23	ABL16645	Drosophila melanog	c 515	16	1.6	129021	21	AAF22296	BAC containing rep
443	16	1.6	6394	24	AAS61278	Human gene regulat	516	16	1.6	145831	24	ABL62309	Colon adenocarcino
444	16	1.6	6394	24	ABK31374	Signal transductio	517	16	1.6	145831	24	ABL66806	Lung cancer relate
445	16	1.6	6651	24	ABL61742	Colon adenocarcino	518	16	1.6	145831	24	ABL68588	Kidney cancer rela
446	16	1.6	6651	24	ABL64868	Lung cancer relate	519	16	1.6	145831	24	ABL69213	Prostate cancer re
447	16	1.6	6651	24	ABL65907	Lung cancer relate	520	16	1.6	147419	24	ABK83574	Human cDNA differe







C 667	15	1.5	456	22	AAK31486	Human bone marrow
C 668	15	1.5	456	24	ABS06240	Human genome-deriv
C 669	15	1.5	459	21	AAC00088	Human secreted pro
C 670	15	1.5	462	22	ABA06980	Human pancreatic c
C 671	15	1.5	462	22	AAK88181	Human digestive sy
C 672	15	1.5	462	23	ABL21521	Drosophila melanog
C 673	15	1.5	466	22	AAK10939	Human breast cance
C 674	15	1.5	467	24	ABK45360	cDNA encoding colo
C 675	15	1.5	470	22	ABA44423	Human breast cell
C 676	15	1.5	470	22	ABA54869	Human foetal liver
C 677	15	1.5	470	22	ABA24635	Probe #3101 for ge
C 678	15	1.5	470	22	AAK03146	Human brain expres
C 679	15	1.5	470	22	AAI13172	Probe #3105 for ge
C 680	15	1.5	470	22	AAI03076	Probe #3067 used t
C 681	15	1.5	475	23	ABK42015	cDNA encoding nove
C 682	15	1.5	477	23	ABV48223	Human prostate exp
C 683	15	1.5	480	23	AAS65295	DNA encoding novel
C 684	15	1.5	481	22	AAK32653	Human bone marrow
C 685	15	1.5	481	22	AAI38468	Probe #7154 used t
C 686	15	1.5	481	24	ABS07444	Human genome-deriv
C 687	15	1.5	481	24	ABQ58368	Human colon cancer
C 688	15	1.5	483	23	ABV13103	Human prostate exp
C 689	15	1.5	485	22	AAI18703	Human breast cance
C 690	15	1.5	486	20	AAK84692	Human metastatic m
C 691	15	1.5	487	24	ABN65261	Human cancer relat
C 692	15	1.5	489	22	AAK18384	Human breast cance
C 693	15	1.5	489	22	AAK81368	Human immune/haema
C 694	15	1.5	490	22	AAK81369	Human immune/haema
C 695	15	1.5	490	23	ABV50595	Human prostate exp
C 696	15	1.5	493	22	AAS46999	Human breast cance
C 697	15	1.5	493	22	AAF17569	Human breast cance
C 698	15	1.5	493	24	ABK95035	Human breast tumou
C 699	15	1.5	495	22	AAI99013	Human excretory re
C 700	15	1.5	495	22	AAI63363	Human kidney relat
C 701	15	1.5	499	24	ABK77477	Bacillus clausii g
C 702	15	1.5	500	24	ABQ39468	Oligonucleotide fo
C 703	15	1.5	500	24	ABQ39469	Oligonucleotide fo
C 704	15	1.5	501	23	AAS86127	DNA encoding novel
C 705	15	1.5	504	21	AAC53591	Arabidopsis thalia
C 706	15	1.5	505	21	AAF08711	Fusarium venenatum
C 707	15	1.5	511	21	AAC52449	Arabidopsis thalia
C 708	15	1.5	511	22	AAI19732	Human breast cance
C 709	15	1.5	514	22	ABA61776	Human foetal liver
C 710	15	1.5	514	22	ABA29382	Probe #7848 for ge
C 711	15	1.5	514	22	AAK10088	Human brain expres
C 712	15	1.5	514	22	AAK35982	Human bone marrow
C 713	15	1.5	514	22	AAI17195	Probe #7128 for ge
C 714	15	1.5	514	22	AAI41696	Probe #10382 used
C 715	15	1.5	514	24	ABS10171	Human genome-deriv
C 716	15	1.5	518	22	AAK61942	Oligonucleotide fo
C 717	15	1.5	520	24	ABQ25322	Oligonucleotide fo
C 718	15	1.5	520	24	ABQ25323	Oligonucleotide fo
C 719	15	1.5	521	22	AAH81622	Human differential
C 720	15	1.5	521	24	ABQ24696	Oligonucleotide fo
C 721	15	1.5	521	24	ABQ24697	Oligonucleotide fo
C 722	15	1.5	521	24	ABQ30308	Oligonucleotide fo
C 723	15	1.5	521	24	ABQ30309	Oligonucleotide fo
C 724	15	1.5	522	24	ABQ59375	Human colon cancer
C 725	15	1.5	528	24	ABQ29880	Oligonucleotide fo
C 726	15	1.5	528	24	ABQ29881	Oligonucleotide fo
C 727	15	1.5	530	20	AAK20913	Polynucleotide seq
C 728	15	1.5	533	23	ABV50889	Human prostate exp
C 729	15	1.5	534	24	ABN73537	Bovine embryonic g
C 730	15	1.5	535	21	AAZ80214	Human colon cancer
C 731	15	1.5	536	23	ABV01367	Human prostate exp
C 732	15	1.5	538	24	ABQ38002	Oligonucleotide fo
C 733	15	1.5	538	24	ABQ38003	Oligonucleotide fo
C 734	15	1.5	542	24	ABK45851	cDNA encoding colo
C 735	15	1.5	544	24	ABQ36756	Oligonucleotide fo
C 736	15	1.5	544	24	ABQ36757	Oligonucleotide fo
C 737	15	1.5	548	22	ABA59546	Human foetal liver
C 738	15	1.5	548	22	AAK07819	Human brain expres
C 739	15	1.5	548	22	AAK33685	Human bone marrow
C 740	15	1.5	740	15	C 740	Human bone marrow
C 741	15	1.5	741	15	C 741	Human genome-deriv
C 742	15	1.5	742	15	C 742	Human secreted pro
C 743	15	1.5	743	15	C 743	Human pancreatic c
C 744	15	1.5	744	15	C 744	Human digestive sy
C 745	15	1.5	745	15	C 745	Drosophila melanog
C 746	15	1.5	746	15	C 746	Human breast cance
C 747	15	1.5	747	15	C 747	cDNA encoding colo
C 748	15	1.5	748	15	C 748	Human breast cell
C 749	15	1.5	749	15	C 749	Human foetal liver
C 750	15	1.5	750	15	C 750	Probe #3101 for ge
C 751	15	1.5	751	15	C 751	Human brain expres
C 752	15	1.5	752	15	C 752	Probe #3105 for ge
C 753	15	1.5	753	15	C 753	Probe #3067 used t
C 754	15	1.5	754	15	C 754	cDNA encoding nove
C 755	15	1.5	755	15	C 755	Human prostate exp
C 756	15	1.5	756	15	C 756	DNA encoding novel
C 757	15	1.5	757	15	C 757	Human bone marrow
C 758	15	1.5	758	15	C 758	Probe #7154 used t
C 759	15	1.5	759	15	C 759	Human genome-deriv
C 760	15	1.5	760	15	C 760	Human colon cancer
C 761	15	1.5	761	15	C 761	Human prostate exp
C 762	15	1.5	762	15	C 762	Human breast cance
C 763	15	1.5	763	15	C 763	Human metastatic m
C 764	15	1.5	764	15	C 764	Human cancer relat
C 765	15	1.5	765	15	C 765	Human breast cance
C 766	15	1.5	766	15	C 766	Human immune/haema
C 767	15	1.5	767	15	C 767	Human immune/haema
C 768	15	1.5	768	15	C 768	Human prostate exp
C 769	15	1.5	769	15	C 769	Human breast cance
C 770	15	1.5	770	15	C 770	Human breast cance
C 771	15	1.5	771	15	C 771	Human breast tumou
C 772	15	1.5	772	15	C 772	Human excretory re
C 773	15	1.5	773	15	C 773	Human kidney relat
C 774	15	1.5	774	15	C 774	Bacillus clausii g
C 775	15	1.5	775	15	C 775	Oligonucleotide fo
C 776	15	1.5	776	15	C 776	Oligonucleotide fo
C 777	15	1.5	777	15	C 777	DNA encoding novel
C 778	15	1.5	778	15	C 778	Arabidopsis thalia
C 779	15	1.5	779	15	C 779	Fusarium venenatum
C 780	15	1.5	780	15	C 780	Arabidopsis thalia
C 781	15	1.5	781	15	C 781	Human breast cance
C 782	15	1.5	782	15	C 782	Human foetal liver
C 783	15	1.5	783	15	C 783	Probe #7848 for ge
C 784	15	1.5	784	15	C 784	Human brain expres
C 785	15	1.5	785	15	C 785	Human bone marrow
C 786	15	1.5	786	15	C 786	Probe #7128 for ge
C 787	15	1.5	787	15	C 787	Probe #10382 used
C 788	15	1.5	788	15	C 788	Human genome-deriv
C 789	15	1.5	789	15	C 789	Human immune/haema
C 790	15	1.5	790	15	C 790	Oligonucleotide fo
C 791	15	1.5	791	15	C 791	Oligonucleotide fo
C 792	15	1.5	792	15	C 792	Human differential
C 793	15	1.5	793	15	C 793	Oligonucleotide fo
C 794	15	1.5	794	15	C 794	Oligonucleotide fo
C 795	15	1.5	795	15	C 795	Oligonucleotide fo
C 796	15	1.5	796	15	C 796	Oligonucleotide fo
C 797	15	1.5	797	15	C 797	Human colon cancer
C 798	15	1.5	798	15	C 798	Oligonucleotide fo
C 799	15	1.5	799	15	C 799	Oligonucleotide fo
C 800	15	1.5	800	15	C 800	Polynucleotide seq
C 801	15	1.5	801	15	C 801	Human prostate exp
C 802	15	1.5	802	15	C 802	Bovine embryonic g
C 803	15	1.5	803	15	C 803	Human colon cancer
C 804	15	1.5	804	15	C 804	Human prostate exp
C 805	15	1.5	805	15	C 805	Oligonucleotide fo
C 806	15	1.5	806	15	C 806	Oligonucleotide fo
C 807	15	1.5	807	15	C 807	cDNA encoding colo
C 808	15	1.5	808	15	C 808	Oligonucleotide fo
C 809	15	1.5	809	15	C 809	Oligonucleotide fo
C 810	15	1.5	810	15	C 810	Human foetal liver
C 811	15	1.5	811	15	C 811	Human brain expres
C 812	15	1.5	812	15	C 812	Human bone marrow
C 813	15	1.5	813	15	C 813	Human bone marrow
C 814	15	1.5	814	15	C 814	Human genome-deriv
C 815	15	1.5	815	15	C 815	Human secreted pro
C 816	15	1.5	816	15	C 816	Human pancreatic c
C 817	15	1.5	817	15	C 817	Human digestive sy
C 818	15	1.5	818	15	C 818	Drosophila melanog
C 819	15	1.5	819	15	C 819	Human breast cance
C 820	15	1.5	820	15	C 820	cDNA encoding colo
C 821	15	1.5	821	15	C 821	Human breast cell
C 822	15	1.5	822	15	C 822	Human foetal liver
C 823	15	1.5	823	15	C 823	Probe #3101 for ge
C 824	15	1.5	824	15	C 824	Human brain expres
C 825	15	1.5	825	15	C 825	Probe #3105 for ge
C 826	15	1.5	826	15	C 826	Probe #3067 used t
C 827	15	1.5	827	15	C 827	cDNA encoding nove
C 828	15	1.5	828	15	C 828	Human prostate exp
C 829	15	1.5	829	15	C 829	DNA encoding novel
C 830	15	1.5	830	15	C 830	Human bone marrow
C 831	15	1.5	831	15	C 831	Probe #7154 used t
C 832	15	1.5	832	15	C 832	Human genome-deriv
C 833	15	1.5	833	15	C 833	Human colon cancer
C 834	15	1.5	834	15	C 834	Human prostate exp
C 835	15	1.5	835	15	C 835	Human breast cance
C 836	15	1.5	836	15	C 836	Human immune/haema
C 837	15	1.5	837	15	C 837	Human immune/haema
C 838	15	1.5	838	15	C 838	Human prostate exp
C 839	15	1.5	839	15	C 839	Human breast cance
C 840	15	1.5	840	15	C 840	Human breast cance
C 841	15	1.5	841	15	C 841	Human breast tumou
C 842	15	1.5	842	15	C 842	Human excretory re
C 843	15	1.5	843	15	C 843	Human kidney relat
C 844	15	1.5	844	15	C 844	Bacillus clausii g
C 845	15	1.5	845	15	C 845	Oligonucleotide fo
C 846	15	1.5	846	15	C 846	Oligonucleotide fo
C 847	15	1.5	847	15	C 847	DNA encoding novel
C 848	15	1.5	848	15	C 848	Arabidopsis thalia
C 849	15	1.5	849	15	C 849	Fusarium venenatum
C 850	15	1.5	850	15	C 850	Arabidopsis thalia
C 851	15	1.5	851	15	C 851	Human breast cance
C 852	15	1.5	852	15	C 852	Human foetal liver
C 853	15	1.5	853	15	C 853	Probe #7848 for ge
C 854	15	1.5	854	15	C 854	Human brain expres
C 855	15	1.5	855	15	C 855	Human bone marrow
C 856	15	1.5	856	15	C 856	Probe #7128 for ge
C 857	15	1.5	857	15	C 857	Probe #10382 used
C 858	15	1.5	858	15	C 858	Human genome-deriv
C 859	15	1.5</				

C 813	15	1.5	685	22	AAH05052	Human cDNA clone (	886	15	1.5	926	21	AAF08966	Fusarium venenatum
C 814	15	1.5	690	24	ABK27480	DNA encoding novel	887	15	1.5	929	24	ABL60676	Z. dhumnades SCAR
815	15	1.5	693	22	AAH67481	C glutamicum codin	C 888	15	1.5	930	22	AAH53107	S. epidermidis ope
816	15	1.5	695	24	AAS62895	Human G protein-co	C 889	15	1.5	941	20	AAH51718	DNA encoding a hum
C 817	15	1.5	700	22	AAH93088	Human inflammatory	890	15	1.5	942	22	AAH68175	C glutamicum codin
C 818	15	1.5	700	22	AAH93089	Human inflammatory	891	15	1.5	942	24	ABN69037	Streptococcus poly
C 819	15	1.5	705	22	AAH08658	Human cDNA clone (	892	15	1.5	943	22	AAK88548	Human digestive sy
820	15	1.5	719	21	AAF12103	Aspergillus oryzae	893	15	1.5	945	20	AAZ09139	B. flavum lactate
821	15	1.5	720	24	AAS62104	Porcine muscular s	C 894	15	1.5	946	19	AAV44192	Lettuce resistance
C 822	15	1.5	732	22	AAL19564	Human breast cance	C 895	15	1.5	946	24	ABK67789	Lettuce pest resis
C 823	15	1.5	733	18	AAV01396	Bovine Factor XIII	C 896	15	1.5	951	17	AAT13641	ACNPV ORF 25, resi
C 824	15	1.5	734	24	AAH07246	Human cDNA clone (	897	15	1.5	970	23	AAS92357	DNA encoding novel
825	15	1.5	744	24	ABQ16942	Oligonucleotide fo	898	15	1.5	987	21	AAA95473	E. coli essential
C 826	15	1.5	744	24	ABQ16943	Oligonucleotide fo	899	15	1.5	987	21	AAA88700	E. coli FUN essent
C 827	15	1.5	749	23	AAS74225	DNA encoding novel	C 900	15	1.5	992	24	ABK74425	Bacillus lichenifo
828	15	1.5	764	22	AAI01238	GIPI2-C4 coding se	901	15	1.5	995	24	ABQ15524	Oligonucleotide fo
C 829	15	1.5	766	22	ABA47563	Human breast cell	C 902	15	1.5	995	24	ABQ15525	Oligonucleotide fo
C 830	15	1.5	766	22	ABA65456	Human foetal liver	C 903	15	1.5	996	24	AAL44645	Protein phosphatas
C 831	15	1.5	766	22	ABA32553	Probe #11019 for g	C 904	15	1.5	998	22	AAH34507	Human colon cancer
C 832	15	1.5	766	22	AAK13864	Human brain expres	905	15	1.5	998	24	ABL90335	Human polynucleoti
C 833	15	1.5	766	22	AAK39608	Human bone marrow	906	15	1.5	1001	21	AAC57749	Arachidonic acid m
C 834	15	1.5	766	22	AAI20418	Probe #10351 for g	907	15	1.5	1001	21	AAC57750	Arachidonic acid m
C 835	15	1.5	766	22	AAI45622	Probe #14308 used	908	15	1.5	1001	21	AAC57751	Arachidonic acid m
C 836	15	1.5	766	22	AAI06118	Probe #6109 used t	909	15	1.5	1001	21	AAC57752	Arachidonic acid m
C 837	15	1.5	766	24	ABS13699	Human genome-deriv	910	15	1.5	1001	22	AAF91430	Moraxella catarrha
838	15	1.5	767	22	AAH08513	Human cDNA clone (	911	15	1.5	1001	24	ABK37810	DNA sequence upstr
C 839	15	1.5	768	23	ABV10536	Human prostate exp	912	15	1.5	1017	24	ABQ54520	Human ovarian anti
840	15	1.5	770	24	ABK74766	Bacillus lichenifo	913	15	1.5	1020	20	AAZ40840	Secreted protein E
C 841	15	1.5	771	22	ABA18829	Human nervous syst	914	15	1.5	1021	24	ABL64876	Lung cancer relate
C 842	15	1.5	772	22	AAI96386	Human neuroblastom	915	15	1.5	1021	24	ABL65631	Lung cancer relate
C 843	15	1.5	772	24	AAD23979	Beta-2 adrenergic	916	15	1.5	1033	23	ABL04021	Drosophila melanog
C 844	15	1.5	778	20	AAZ16583	Human gene express	C 917	15	1.5	1036	19	AAV44191	Lettuce resistance
845	15	1.5	778	20	AAX37416	Human secreted pro	C 918	15	1.5	1036	24	ABK67788	Lettuce pest resis
846	15	1.5	782	21	AAA01907	Human colon cancer	919	15	1.5	1040	21	AAC46418	Arabidopsis thalia
C 847	15	1.5	784	22	AAI97246	Human neuroblastom	920	15	1.5	1042	21	AAC37336	Arabidopsis thalia
C 848	15	1.5	789	18	AAT79323	DNA encoding Archa	921	15	1.5	1050	21	AAA50033	DNA encoding Chlam
849	15	1.5	789	20	AAZ16233	Human gene express	922	15	1.5	1065	22	AAF71424	Corynebacterium gl
850	15	1.5	789	22	AAH07981	Human cDNA clone (	923	15	1.5	1069	24	ABQ68667	Listeria monocytog
C 851	15	1.5	795	23	AAS53887	Helicobacter pylor	924	15	1.5	1074	23	ABL29061	Drosophila melanog
C 852	15	1.5	799	22	AAH01667	Shewanella putrefa	C 925	15	1.5	1074	24	ABN91079	Staphylococcus epi
C 853	15	1.5	804	19	AAV28698	Ripening banana pu	926	15	1.5	1076	22	AAI59902	Human polynucleoti
854	15	1.5	804	24	ABN79898	Fungal ZBC gene se	C 927	15	1.5	1078	22	AAD13879	Human PUMP-1 anti
C 855	15	1.5	807	22	AAH53218	S. epidermidis ope	928	15	1.5	1078	22	AAD13880	Human PUMP-1 cDNA
856	15	1.5	810	19	AAV28665	Ripening banana pu	929	15	1.5	1078	22	AAH28224	Nucleotide sequenc
857	15	1.5	810	22	AAS42027	Genomic sequence #	C 930	15	1.5	1078	24	ABK44207	CDNA #147 encoding
C 858	15	1.5	812	22	AAH70190	Human cervical can	C 931	15	1.5	1093	24	AAD23980	Beta-2 adrenergic
859	15	1.5	813	23	AAS88501	DNA encoding novel	932	15	1.5	1094	20	AAX13577	Enterococcus faeca
860	15	1.5	816	20	AAX36834	Human XLIS gene fr	933	15	1.5	1096	24	ABK75102	Bacillus lichenifo
861	15	1.5	820	24	ABQ89121	Human prostate exp	934	15	1.5	1110	21	AAC79664	Virulence gene #71
C 862	15	1.5	821	22	AAF56508	Zea mays root cap	935	15	1.5	1124	22	AAS33253	DNA encoding human
C 863	15	1.5	822	22	AAH05403	Human cDNA clone (	936	15	1.5	1125	21	AAA75886	CDNA encoding a hu
C 864	15	1.5	822	24	ABN91617	Staphylococcus epi	937	15	1.5	1125	22	AAH74999	Nucleotide sequenc
865	15	1.5	829	24	ABA94544	NDV F protein epit	938	15	1.5	1135	22	AAK71053	Human immune/haema
C 866	15	1.5	832	22	AAK92300	Human cDNA 5'-end	939	15	1.5	1137	21	AAC47343	Arabidopsis thalia
C 867	15	1.5	832	22	AAK93824	Human cDNA clone r	C 940	15	1.5	1137	24	ABK75018	Bacillus lichenifo
868	15	1.5	835	23	AAS90047	DNA encoding novel	941	15	1.5	1142	24	ABQ72492	Human MDDT encodin
869	15	1.5	842	22	AAH06424	Human cDNA clone (	942	15	1.5	1143	21	AAC38217	Arabidopsis thalia
C 870	15	1.5	843	22	AAK75591	Human immune/haema	C 943	15	1.5	1143	24	AAS20573	Human uterine glob
C 871	15	1.5	843	22	AAK75592	Human immune/haema	C 944	15	1.5	1144	24	ABK73513	Bacillus lichenifo
C 872	15	1.5	846	23	ABV03404	Human prostate exp	945	15	1.5	1157	22	AAF98704	Human ovarian canc
C 873	15	1.5	849	23	AAS72804	DNA encoding novel	946	15	1.5	1167	22	AAH68007	C glutamicum codin
874	15	1.5	855	21	AAAS9477	Nucleotide sequenc	947	15	1.5	1173	22	AAH35031	Human colon cancer
C 875	15	1.5	875	21	AAL12064	Human breast cance	948	15	1.5	1175	23	AAS85060	DNA encoding novel
876	15	1.5	879	22	AAF27604	Mevalonate pathway	C 949	15	1.5	1177	24	AAS16003	DNA encoding human
C 877	15	1.5	885	22	AAH04976	Human cDNA clone (	950	15	1.5	1184	24	ABK71705	Human dithp polynu
C 878	15	1.5	885	23	ABL25057	Drosophila melanog	C 951	15	1.5	1185	23	ABL11145	Drosophila melanog
879	15	1.5	900	22	AAH65608	C glutamicum codin	C 952	15	1.5	1186	21	AAF12872	Aspergillus oryzae
880	15	1.5	902	22	AAK90705	Human digestive sy	953	15	1.5	1187	21	AAC53579	Arabidopsis thalia
C 881	15	1.5	908	20	AAX37412	Human secreted pro	954	15	1.5	1191	24	ABQ54821	Human ovarian anti
C 882	15	1.5	908	20	AAX07168	Corn dihydrodipico	C 955	15	1.5	1203	21	AAC95498	Human secreted pro
C 883	15	1.5	912	21	AAC51606	Arabidopsis thalia	C 956	15	1.5	1214	22	AAF59786	Arabidopsis thalia
C 884	15	1.5	914	21	AAC33369	Arabidopsis thalia	957	15	1.5	1217	20	AAX20432	Human secreted pro
885	15	1.5	921	24	ABN69078	Streptococcus poly	C 958	15	1.5	1230	24	ABN68925	Streptococcus poly

c 959 15 1.5 1242 21 AAA62991  
960 15 1.5 1242 23 AAS71485  
961 15 1.5 1248 23 ABL25803  
962 15 1.5 1248 24 AAD21998  
963 15 1.5 1255 23 AAS86210  
964 15 1.5 1258 24 AAS94877  
965 15 1.5 1262 24 AAD36247  
966 15 1.5 1266 21 AAC51603  
967 15 1.5 1266 22 AAH52375  
968 15 1.5 1267 21 AAC77794  
969 15 1.5 1287 12 AAQ13381  
970 15 1.5 1300 20 AAZ09352  
971 15 1.5 1302 24 ABK75350  
972 15 1.5 1305 21 AAC54212  
973 15 1.5 1305 23 AAS89621  
974 15 1.5 1310 19 AAV52399  
975 15 1.5 1311 24 ABN90923  
976 15 1.5 1319 21 AAA38558  
977 15 1.5 1320 21 AAC43225  
978 15 1.5 1330 21 AAC41175  
979 15 1.5 1330 22 AAH23816  
980 15 1.5 1335 24 ABN66154  
981 15 1.5 1340 21 AAC69477  
982 15 1.5 1346 24 ABN81333  
983 15 1.5 1347 22 AAF60981  
984 15 1.5 1353 11 AAQ05667  
985 15 1.5 1353 11 AAQ05759  
986 15 1.5 1353 11 AAQ08572  
987 15 1.5 1353 11 AAQ08544  
988 15 1.5 1353 18 AAT60952  
989 15 1.5 1355 21 AAC46835  
990 15 1.5 1356 23 ABL27499  
991 15 1.5 1358 24 AAS62425  
992 15 1.5 1370 21 AAA26413  
993 15 1.5 1374 22 AAS40754  
994 15 1.5 1374 22 AAL07400  
995 15 1.5 1375 22 AAS40755  
996 15 1.5 1375 22 AAS40756  
997 15 1.5 1375 22 AAL07401  
998 15 1.5 1375 22 AAL07402  
999 15 1.5 1379 23 AAS72805  
1000 15 1.5 1382 23 ABV25632

ALIGNMENTS

RESULT 1  
AAN80494  
ID AAN80494 standard; DNA; 1234 BP.  
XX  
AC AAN80494;  
XX  
DT 29-NOV-1990 (first entry)  
XX  
DE Sequence encoding human testis-specific lactate dehydrogenase subunit.  
XX  
KW Testis-specific lactate dehydrogenase; vaccine; female fertility; ss.  
XX  
OS synthetic.  
XX  
FH Key Location/Qualifiers  
FT CDS 76..1083  
FT /\*tag= a  
FT /label=LDH-C4  
XX  
PN EP270056-A.  
XX  
PD 08-JUN-1988.  
XX  
PF 30-NOV-1987; 87EP-0117688.  
XX  
PR 12-JUN-1987; 87US-0061894.

Human beta2-adrene  
DNA encoding novel  
Drosophila melanog  
Human transporters  
DNA encoding novel  
Human DNA sequence  
Human TWIK-9 cDNA.  
Arabidopsis thalia  
S. epidermidis ope  
Human cancer assoc  
Human beta2-adrene  
Plasmid PUSlim DNA  
Bacillus lichenifo  
Arabidopsis thalia  
DNA encoding novel  
Streptococcus pneu  
Staphylococcus epi  
Actinobacillus ple  
Arabidopsis thalia  
Arabidopsis thalia  
Human transferase  
Streptococcus poly  
Human secreted pro  
Physcomitrella pat  
P. putida KT2440-a  
pMS10 contg. male  
Clone pMS10 contg.  
cDNA clone pMS10 e  
cDNA clone pMS10 e  
Anther-specific cD  
Arabidopsis thalia  
Drosophila melanog  
cDNA sequence #212  
Human secreted pro  
DNA encoding human  
Human reproductive  
DNA encoding human  
DNA encoding human  
Human reproductive  
Human reproductive  
DNA encoding novel  
Human prostate exp

PR 12-JUN-1987; 87US-0936451.  
XX  
PA (NOUN ) NORTHWESTERN UNIV.  
XX  
PI Goldberg E, Millan JL;  
XX  
DR WPI; 1988-156290/23.  
DR P-PSDB; AAP80891.  
XX  
PT New recombinant DNA encoding sub-unit of testis lactate hydrogenase -  
PT expressing antigenic polypeptide useful in vaccines for reducing  
PT female fertility.  
XX  
PS Disclosure; ; pp; English.  
XX  
CC This encodes a testis-specific lactate dehydrogenase sub-unit (LDH-C4).  
CC The complete subunit or one or more antigenic regions can be used in a  
CC vaccine prepn. to reduce female fertility. Larger peptides (or fusion  
CC proteins) expressed by this DNA are antigenic and smaller peptides be-  
CC come so when coupled to e.g. tetanus toxoid.  
XX  
SQ Sequence 1234 BP; 366 A; 202 C; 286 G; 380 T; 0 other;  
  
Query Match 2.1%; Score 20; DB 9; Length 1234;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 547 ATTGGAGAACATGGTGATTC 566  
|||||||  
Db 643 ATTGGAGAACATGGTGATTC 662  
  
RESULT 2  
AAS17494  
ID AAS17494 standard; cDNA; 1680 BP.  
XX  
AC AAS17494;  
XX  
DT 26-FEB-2002 (first entry)  
XX  
DE Human cDNA encoding testicular lactate dehydrogenase A.  
XX  
KW Human; ss; testicular lactate dehydrogenase A; fusion protein;  
KW gene expression chip.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 99..1244  
FT /\*tag= a  
FT /product= "Testicular lactate dehydrogenase A"  
XX  
PN CN1313342-A.  
XX  
PD 19-SEP-2001.  
XX  
PF 28-FEB-2001; 2001CN-0108262.  
XX  
PR 28-FEB-2001; 2001CN-0108262.  
XX  
PA (UYNA-) UNIV NANJING MEDICAL.  
XX  
PI Sha J, Zhou Z, Li J;  
XX  
DR WPI; 2002-042206/06.  
DR P-PSDB; AAU11432.  
XX  
PT New human testis lactate dehydrogenase A for preparing monoclonal and  
PT multiclonal antibodies -  
XX  
PS Claim 4; Page 1-2; 6pp; Chinese.  
XX  
CC The invention relates to a human testicular novel lactate dehydrogenase

CC A with 1680 base pairs (bp) of cDNA sequence, 1146 bp of open reading  
CC frame, 382 coding amino acids and number AY009108 in Genbank.  
CC The cDNA can be used to prepare a fusion protein and a  
CC gene expression chip of testicular specific function. The fusion  
CC protein can be used to prepare monoclonal and multiclional antibodies.  
CC The present sequence is the cDNA encoding testicular lactate  
CC dehydrogenase A.  
XX  
SQ Sequence 1680 BP; 489 A; 335 C; 368 G; 488 T; 0 other;  
  
Query Match 2.0%; Score 19; DB 24; Length 1680;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 471 TCTGGATACTGCTCGTTT 489  
Db 737 TCTGGATACTGCTCGTTT 755  
  
RESULT 3  
AAH64739  
ID AAH64739 standard; cDNA; 1755 BP.  
XX  
AC AAH64739;  
XX 11-SEP-2001 (first entry)  
XX Human secreted protein cDNA, SEQ ID NO: 15.  
DE Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;  
KW GENSET; ss.  
XX Homo sapiens.  
OS  
XX WO200142451-A2.  
PN  
XX 14-JUN-2001.  
PD  
XX  
PF 07-DEC-2000; 2000WO-IB01938.  
XX  
XX 08-DEC-1999; 99US-0169629.  
PR  
PR 06-MAR-2000; 2000US-0187470.  
XX  
PA (GEST ) GENSET.  
XX  
PI Dumas Milne Edwards J, Bougueleret L, Jobert S;  
XX  
XX WPI; 2001-367870/38.  
DR  
DR P-PSDB; AAG89136.  
XX  
PT Full length GENSET human nucleic acids encoding potentially secreted  
PT proteins, useful in gene therapy and vaccination against a variety of  
PT diseases, and for diagnosis of those diseases -  
XX  
PS Claim 7; Page 574-575; 921pp; English.  
XX  
CC The invention relates to full length GENSET human nucleic acids encoding  
CC potentially secreted proteins. The nucleic acids and the polypeptides  
CC they encode may be used in the prevention, treatment and diagnosis of  
CC diseases associated with inappropriate GENSET gene expression. For  
CC example, they be used to treat disorders associated with decreased  
CC GENSET gene expression by rectifying mutations or deletions in a  
CC patient's genome that affect the activity of GENSET or by supplementing  
CC the patients own production of GENSET polypeptides. Conversely,  
CC antisense nucleic acid molecules may be administered to down regulate  
CC GENSET expression by binding with the cells' own genes and preventing  
CC used as DNA probes in diagnostic assays to detect and quantitate the  
CC presence of similar nucleic acid sequences in samples, and hence to  
CC determine which patients may be in need of restorative therapy.  
CC The GENSET polypeptides may be used as antigens in the production of  
CC antibodies and in assays to identify modulators (agonists and  
CC antagonists) of GENSET polypeptide expression and activity. The

CC present sequence is a GENSET nucleic acid of the invention.  
XX  
SQ Sequence 1755 BP; 514 A; 343 C; 379 G; 519 T; 0 other;  
  
Query Match 2.0%; Score 19; DB 22; Length 1755;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 471 TCTGGATACTGCTCGTTT 489  
Db 737 TCTGGATACTGCTCGTTT 755  
  
RESULT 4  
AAH64738  
ID AAH64738 standard; cDNA; 1759 BP.  
XX  
AC AAH64738;  
XX 11-SEP-2001 (first entry)  
XX Human secreted protein cDNA, SEQ ID NO: 14.  
DE Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;  
KW GENSET; ss.  
XX Homo sapiens.  
OS  
XX WO200142451-A2.  
PN  
XX 14-JUN-2001.  
PD  
XX  
PF 07-DEC-2000; 2000WO-IB01938.  
XX  
XX 08-DEC-1999; 99US-0169629.  
PR  
PR 06-MAR-2000; 2000US-0187470.  
XX  
PA (GEST ) GENSET.  
XX  
PI Dumas Milne Edwards J, Bougueleret L, Jobert S;  
XX  
XX WPI; 2001-367870/38.  
DR  
DR P-PSDB; AAG89135.  
XX  
PT Full length GENSET human nucleic acids encoding potentially secreted  
PT proteins, useful in gene therapy and vaccination against a variety of  
PT diseases, and for diagnosis of those diseases -  
XX  
PS Claim 7; Page 572-574; 921pp; English.  
XX  
CC The invention relates to full length GENSET human nucleic acids encoding  
CC potentially secreted proteins. The nucleic acids and the polypeptides  
CC they encode may be used in the prevention, treatment and diagnosis of  
CC diseases associated with inappropriate GENSET gene expression. For  
CC example, they be used to treat disorders associated with decreased  
CC GENSET gene expression by rectifying mutations or deletions in a  
CC patient's genome that affect the activity of GENSET or by supplementing  
CC the patients own production of GENSET polypeptides. Conversely,  
CC antisense nucleic acid molecules may be administered to down regulate  
CC GENSET expression by binding with the cells' own genes and preventing  
CC their expression. The sense and antisense nucleic acids may also be  
CC used as DNA probes in diagnostic assays to detect and quantitate the  
CC presence of similar nucleic acid sequences in samples, and hence to  
CC determine which patients may be in need of restorative therapy.  
CC The GENSET polypeptides may be used as antigens in the production of  
CC antibodies and in assays to identify modulators (agonists and  
CC antagonists) of GENSET polypeptide expression and activity. The  
XX present sequence is a GENSET nucleic acid of the invention.  
SQ Sequence 1759 BP; 515 A; 344 C; 379 G; 521 T; 0 other;  
  
Query Match 2.0%; Score 19; DB 22; Length 1759;  
Best Local Similarity 100.0%; Pred. No. 18;



Matches	19;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	471	TCTGGATAC	TGCTCGTTTT	489					
Db	739	TCTGGATAC	TGCTCGTTTT	757					
RESULT '5									
AAD14466									
ID	AAD14466	standard;	CDNA; 4068 BP.						
XX	AC	AAD14466;							
XX	01-NOV-2001	(first entry)							
XX	Mouse p97	(mp97) protein	cdna.						
XX	Mouse;	mp97 protein;	scialoglycoprotein;	neuroprotective;	antibacterial;				
KW	analgesic;	nootropic;	cytostatic;	neuroleptic;	virucide;	anticonvulsant;			
KW	deficiency	disease;	Wernicke's	disease;	neurodegenerative	disease;	pain;		
KW	nutritional	polyneuropathy;	neurological	disorder;	cancer;	gene therapy;			
KW	Huntington's	disease;	Alzheimer's	disease;	Parkinson's	disease;	epilepsy;		
KW	demyelinating	disease;	multiple	sclerosis;	amyotrophic	lateral	sclerosis;		
KW	psychosis;	therapeutic;	ss.						
XX	Mus	sp.							
XX	Key	Location/Qualifiers							
FH	CDS	64..2280							
FT		/*tag= a	/*product= "Mouse p97 protein"						
FT		64..120							
FT	sig_peptide	/*tag= b							
FT	mat_peptide	121..2277							
FT		/*tag= c							
FT		/*product= "Mature mouse p97 protein"							
FT	5'UTR	1..63							
FT		/*tag= d							
FT		/*note= "5', untranslated region (UTR)"							
FT	3'UTR	2281..4068							
FT		/*tag= e							
FT		/*note= "3', untranslated region (UTR)"							
FT	polyA_signal	3299..3304							
FT		/*tag= f							
FT		/*note= "Polyadenylation signal I"							
FT	polyA_site	3544							
FT		/*tag= g							
FT		/*note= "Polyadenylation site I"							
FT	polyA_signal	3106..3111							
FT		/*tag= h							
FT		/*note= "Polyadenylation signal II"							
FT	polyA_site	3128							
FT		/*tag= i							
FT		/*note= "Polyadenylation site II"							
FT	polyA_signal	4028..4033							
FT		/*tag= j							
FT		/*note= "Polyadenylation signal for EST2 transcript"							
FT	polyA_site	4048							
FT		/*tag= k							
FT		/*note= "Polyadenylation site for EST2 transcript"							
FT	polyA_site	4049..4068							
FT		/*tag= l							
FT		/*note= "Polyadenylation tail"							
XX	WO200159459-A2.								
PN	16-AUG-2001.								
XX	08-FEB-2001;	2001WO-CA00133.							
XX	08-FEB-2000;	2000US-0181091.							
XX	(UYBR-)	UNIV BRITISH COLUMBIA.							

Cheng N, Gagnier L, Jefferies WA;  
WPI; 2001-514683/56.  
P-PSDB; AAE06668.

Novel murine p97 polypeptides and polynucleotides for preparing experimental models to study murine p97 and to identify modulators of murine p97 expression or activity useful for treating neurological conditions -

Claim 25; Page 53-54; 70pp; English.

The invention relates to mouse p97 protein, mp97 (a scialoglycoprotein) and its corresponding cDNA molecule. Mouse p97 protein and its DNA molecule are useful for identifying compounds that affects mp97 protein activity or expression. The invention also relates to a method for screening therapeutic agents which are useful for treating neurological conditions, such as cancer, neurodegenerative diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease), demyelinating diseases (e.g., multiple sclerosis), amyotrophic lateral sclerosis, bacterial and viral infections, deficiency diseases (e.g., Wernicke's disease, nutritional polyneuropathy), epilepsy, psychosis, pain and neurological disorders, especially Alzheimer's disease. Mouse p97 DNA's are also useful in gene therapy. Mp97 proteins are useful for delivering therapeutic agents and pharmaceuticals across the blood placenta barrier as well as to other organs including liver. The invention is also useful for preparing antibodies and antisense oligonucleotides, the preparation of experimental systems to study mp97, and in diagnostic and therapeutic applications. Transgenic p97 mice is useful for identifying essential physiological roles for p97 in development and adult functioning of the organism and for testing potential therapeutic and diagnostic agents that are conjugated to p97 protein. The present cDNA sequence encodes mouse p97 (mp97) protein.

Sequence 4068 BP; 934 A; 1138 C; 1145 G; 850 T; 1 other;

Query Match 2.0%; Score 19; DB 22; Length 4068;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 803 CCGTTGGTGATTATTTGG 821  
|||||  
Db 551 CCGTTGGTGATTATTTGG 569

RESULT 6  
AAF62198  
ID AAF62198 standard; DNA; 4158 BP.  
XX  
AC AAF62198;  
XX  
DT 21-MAY-2001 (first entry)  
XX  
DE DNA encoding membrane bound transferrin like protein.  
XX  
KW Chondrogenesis promoter; membrane-bound transferrin-like protein; Mtf;  
KW Chondrogenesis regulator; Mtf activator; bone metabolism; mouse;  
KW Chondral differentiation inhibitor; bone disease; ds.  
XX  
OS Mus sp.  
XX  
PN WO200113951-A1.  
XX  
PD 01-MAR-2001.  
XX  
PF 21-AUG-2000; 2000WO-JP05590.  
XX  
PR 19-AUG-1999; 99JP-0232966.  
XX  
PA (CHUS ) CHUGAI SEIYAKU KK.  
XX  
PI Kato Y, Fujimoto K;

```

XX WPI; 2001-218409/22.
DR P-PSDB; AAB62882.
XX
PT Chondrogenesis promoters containing membrane-bound transferrin-like
PT protein, useful in diagnosis, prevention and treatment of diseases due
PT to abnormal chondral metabolism and bone metabolism -
XX
PS Disclosure; Page 43-46; 57pp; Japanese.
XX
CC This invention relates to chondrogenesis promoters containing a
CC membrane-bound transferrin-like protein (MTf). Chondrogenesis promoters,
CC chondrogenesis regulators, Mtf activators, Mtf antagonist-containing
CC chondral differentiation inhibitors are useful in diagnosis, prevention
CC and treatment of diseases due to abnormal chondral metabolism and bone
CC metabolism e.g. bone diseases. The present sequence represents murine DNA
CC encoding Mtf.
XX
SQ Sequence 4158 BP; 960 A; 1159 C; 1177 G; 862 T; 0 other;

  Query Match      2.0%; Score 19; DB 22; Length 4158;
  Best Local Similarity 100.0%; Pred. No. 17;
  Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 803 CCGTTGGTGCATTATTTGG 821
  |||||||
Db 605 CCGTTGGTGCATTATTTGG 623

RESULT 7
ABL28478
ID ABL28478 standard; DNA; 5313 BP.
XX
AC ABL28478;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 36907.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 36907; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
```

```

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5313 BP; 1472 A; 1219 C; 1230 G; 1392 T; 0 other;

  Query Match      2.0%; Score 19; DB 23; Length 5313;
  Best Local Similarity 100.0%; Pred. No. 17;
  Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 GTGCTAGTGGGAGTGGGAA 76
  |||||||
Db 5108 GTGCTAGTGGGAGTGGGAA 5126

RESULT 8
AAF28536
ID AAF28536 standard; DNA; 33140 BP.
XX
AC AAF28536;
XX
DT 04-APR-2001 (first entry)
XX
DE Genomic fragment #23.
XX
KW Genomic library; bacteria; human upper airway; otitis media; sinusitis;
KW bronchopulmonary; endocarditis; meningitis; ss.
XX
OS Moraxella catarrhalis.
XX
PN WO200078968-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16649.
XX
PR 18-JUN-1999; 99US-0140121.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lagace RE, Patterson C, Berg KL;
XX
DR WPI; 2001-041427/05.
XX
PT Genomic library for identifying diagnostic and therapeutic
PT compositions, and for identifying virulence factors, regulatory
PT elements and drug targets, comprises Moraxella catarrhalis nucleic
PT acids -
XX
PS Claim 1; Page 191-199; 545pp; English.
XX
CC The present invention relates to a Moraxella catarrhalis genomic library
CC comprising of a combination of 41 nucleic acid molecules (see
CC AAF28514-AAF28554). The library has a number of uses described in the
CC specification e.g. is useful for identifying diagnostic and therapeutic
CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large
CC aerobic, gram-negative diplococcus, normally found among the bacterial
CC flora of human upper airways. M. catarrhalis is known to cause acute,
CC localised infections such as otitis media, sinusitis and bronchopulmonary
CC infection and life-threatening, systemic diseases including endocarditis
CC and meningitis.
XX
SQ Sequence 33140 BP; 9472 A; 6496 C; 7384 G; 9787 T; 1 other;

  Query Match      2.0%; Score 19; DB 22; Length 33140;
  Best Local Similarity 100.0%; Pred. No. 17;
  Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 908 AAATAGAATTGATGGAAA 926
  |||||||
Db 9745 AAATAGAATTGATGGAAA 9763
```

RESULT 9  
AAF22305  
ID AAF22305 standard; DNA; 1082138 BP.  
XX  
AC AAF22305;  
XX  
DT 20-MAR-2001 (first entry)  
XX  
DE Arabidopsis thaliana chromosome 4 centromere.  
XX  
KW Centromere; michrosome; vector; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200055325-A2.  
XX  
PD 21-SEP-2000.  
XX  
PF 17-MAR-2000; 2000WO-US07392.  
XX  
PR 18-MAR-1999; 99US-0125219.  
PR 01-APR-1999; 99US-0127409.  
PR 18-MAY-1999; 99US-0134770.  
PR 13-SEP-1999; 99US-0153584.  
PR 17-SEP-1999; 99US-0154603.  
XX  
PA (UYCH-) UNIV CHICAGO.  
XX  
PI Preuss D, Copenhaver G, Keith K;  
XX  
DR WPI; 2000-587529/55.  
XX  
PT Recombinant DNA construct comprising a plant centromere, useful for  
PT producing stably inherited michrosomes which can serve as vectors for  
PT the construction of transgenic plant and animal cells  
XX  
PS Claim 68; Page 977-1388; 1449pp; English.  
XX  
CC The present invention relates to a recombinant DNA construct of a plant  
CC (Arabidopsis thaliana) centromere. The constructs are useful for  
CC producing stably inherited michrosomes which can serve as vectors for  
CC the construction of transgenic plant and animal cells expressing  
CC selected proteins such as hormones, enzymes, interleukins, clotting  
CC factors, cytokines, antibodies, and growth factors.  
XX  
SQ Sequence 1082138 BP; 348775 A; 194404 C; 195515 G; 343444 T; 0 other;  
  
Query Match 2.0%; Score 19; DB 21; Length 1082138;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 908 AAATAGAAATTGATGGAAAA 926  
Db 451974 AAATAGAAATTGATGGAAAA 451992  
  
RESULT 10  
ABL19001/c  
ID ABL19001 standard; DNA; 360 BP.  
XX  
AC ABL19001;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 8476.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX

PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions  
XX  
PS Claim 1; SEQ ID NO 8476; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 360 BP; 105 A; 76 C; 101 G; 78 T; 0 other;  
  
Query Match 1.9%; Score 18; DB 23; Length 360;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 468 AGTCTGGGATACTGCTCG 485  
Db 107 AGTCTGGGATACTGCTCG 90  
  
RESULT 11  
AAF07721/c  
ID AAF07721 standard; cDNA; 644 BP.  
XX  
AC AAF07721;  
XX  
DT 13-MAR-2001 (first entry)  
XX  
DE Fusarium venenatum EST SEQ ID NO:244.  
XX  
KW Multiple gene expression; filamentous fungal cell; EST;  
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;  
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
KW culture condition; environmental stress; spore morphogenesis;  
KW metabolic pathway engineering; catabolic pathway engineering; ss.  
XX  
OS Fusarium venenatum.  
XX  
PN WO200056762-A2.  
XX  
PD 28-SEP-2000.  
XX  
PF 22-MAR-2000; 2000WO-US07781.  
XX  
PR 22-MAR-1999; 99US-0273623.  
XX  
PA (NOVO ) NOVO NORDISK BIOTECH INC.  
PA (NOVO ) NOVO NORDISK AS.  
XX  
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
XX

DR WPI; 2000-594572/56.

XX Monitoring differential expression of genes in filamentous fungal cells

PT uses fluorescence-labeled nucleic acids isolated from the cells and a

PT substrate of expressed sequence tags -

XX

PS Claim 86; Page 476; 3161pp; English.

XX

CC The present invention describes a method for monitoring differential

CC expression of genes in a first filamentous fungal (FF) cell relative to

CC expression of the same genes in one or more second filamentous fungal

CC cells. The method uses fluorescence-labeled nucleic acids isolated from

CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs

CC are used in the methods for monitoring differential expression of genes

CC in a first filamentous fungal (FF) cell relative to expression of the

CC same genes in one or more second filamentous fungal cells. Monitoring

CC the global expression of genes from FF cells allows the production

CC potential of the microorganisms to be improved. New genes may be

CC discovered, possible functions of unknown open reading frames can be

CC identified and gene copy number variation and stability can be

CC monitored. The expression of genes can be used to study how FF cells

CC adapt to changes in culture conditions, environmental stress, spore

CC morphogenesis, recombination, metabolic or catabolic pathway

CC engineering. Using ESTs provides several advantages over genomic or

CC random cDNA clones including elimination of redundancy as one spot on an

CC array equals one gene or open reading frame, and organisation of the

CC microarrays based on function of the gene products to facilitate

CC analysis of the results. AAF07478 to AAF11247 represents ESTs from

CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus

CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and

CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are

CC all specifically claimed in the present invention.

XX

SQ Sequence 644 BP; 174 A; 190 C; 140 G; 132 T; 8 other;

Query Match 1.9%; Score 18; DB 21; Length 644;

Best Local Similarity 100.0%; Pred. No. 60;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 805 GTTGGTGATATTATTGGG 822

Db 23 GTTGGTGATATTATTGGG 6

RESULT 12

ABK79311/c

ID ABK79311 standard; DNA; 828 BP.

XX

AC ABK79311;

XX

DT 13-AUG-2002 (first entry)

XX

DE Bacillus clausii genomic sequence tag (GST) #2154.

XX

KW Differential gene expression; genomic sequenced tag; GST;

KW altered culture condition; environmental stress;

KW physiological provocation; ds.

XX

OS Bacillus clausii.

XX

PN WO200229113-A2.

XX

PD 11-APR-2002.

XX

PF 05-OCT-2001; 2001WO-US31437.

XX

PR 06-OCT-2000; 2000US-0680598.

PR 27-MAR-2001; 2001US-279526P.

XX

PA (NOVO ) NOVOZYMES BIOTECH INC.

PA (NOVO ) NOVOZYMES AS.

XX

PI Berka R, Clausen IG;

XX WPI; 2002-416684/44.

DR

XX Monitoring differential expression of several genes in first Bacillus

PT cell relative to expression of same genes in one or more second

PT Bacillus cells, by using substrate containing Bacillus genomic

PT sequenced tag array -

XX

PS Claim 11; SEQ ID NO 6602; 200pp; English.

XX

CC The invention describes a method of monitoring differential expression of

CC genes in a first Bacillus cell relative to expression of the genes in

CC other Bacillus cells, comprising hybridising labelled nucleic acid probes

CC isolated from Bacillus cells to a substrate containing array of Bacillus

CC genomic sequenced tags (GST), examining the array, and determining

CC relative gene expression by an observed hybridisation reporter signal of

CC a spot in the array. The method is useful for measuring the expression of

CC genes in a first Bacillus cell relative to expression of the same genes

CC in one or more second Bacillus cells. The method is useful for monitoring

CC global expression of several genes from a Bacillus cell, discovering new

CC genes, identifying possible functions of unknown open reading frames and

CC monitoring gene copy number variation and stability. Monitoring changes

CC in expression of genes may be used to provide a representation of the way

CC in which Bacillus cells adapt to changes in culture conditions,

CC environmental stress or other physiological provocation. Extensive

CC follow-up characterisation is unnecessary, when one spot on an array

CC equals one gene or one open reading frame, since sequence information is

CC available. This sequence represents a genomic sequence tag (GST) used in

CC the method of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at

CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 828 BP; 277 A; 136 C; 186 G; 229 T; 0 other;

Query Match 1.9%; Score 18; DB 24; Length 828;

Best Local Similarity 100.0%; Pred. No. 60;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 346 AATGTTGCCAAATATGCT 363

Db 291 AATGTTGCCAAATATGCT 274

RESULT 13

AAS90935/c

ID AAS90935 standard; cDNA; 1393 BP.

XX

AC AAS90935;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #26739.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US08631.

XX

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;



```
DR WPI; 2001-639362/73.
DR P-PSDB; ABG26748.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID No 26739; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1393 BP; 352 A; 352 C; 356 G; 333 T; 0 other;

Query Match 1.9%; Score 18; DB 23; Length 1393;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TCCAAGATACAAAGTCTC 22
Db 146 TCCAAGATACAAAGTCTC 129

RESULT 14
AAZ37126
ID AAZ37126 standard; DNA; 1728 BP.
XX
AC AAZ37126;
XX
DT 28-JAN-2000 (first entry)
XX
DE Nucleotide sequence of H.influenzae HI1655 gene.
XX
KW Genome; mutagenesis; transposon; isolate; locate; essential gene; ss;
KW detect; growth; anti-microbial therapy; genomic footprinting.
XX
OS Haemophilus influenzae.
XX
PN WO9950402-A1.
XX
PD 07-OCT-1999.
XX
PF 26-MAR-1999; 99WO-US06139.
XX
PR 27-MAR-1998; 98US-0079770.
XX
PA (HARD ) HARVARD COLLEGE.
XX
PI Mekalanos JJ, Akerley B, Rubin E, Camilli A;
XX
DR WPI; 1999-620065/53.
XX
PT Detection of genes essential for the growth and viability of organism,
```

```
PT useful as e.g. targets for anti-microbial therapy
XX
PS Disclosure; Fig 5N; 62pp; English.
XX
CC Nucleotide sequences AAZ37113-Z37126 are essential Haemophilus
CC influenzae genes that have been identified by the methods of the
CC invention. The invention relates to a method for locating essential
CC regions of a portion of an organism's genome by: (i) in vitro
CC mutagenising DNA having the sequence of the region with a transposon;
CC (ii) identifying cells of the organism transformed with the mutagenised
CC DNA of (i); and (iii) locating the region by detecting the absence of
CC transposons in the region of the mutagenised cells containing the
CC mutagenised DNA. The invention also relates to a method for isolating a
CC compound that modulates the expression of a nucleic acid sequence
CC operably linked to a gene promoter, and a method of identifying a nucleic
CC acid sequence that is essential for cell growth or viability. The methods
CC are used to detect genes that are essential for the growth and viability
CC of organism. Such genes can be used e.g. as targets for anti-microbial
CC therapy.
XX
SQ Sequence 1728 BP; 574 A; 332 C; 343 G; 479 T; 0 other;

Query Match 1.9%; Score 18; DB 20; Length 1728;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 TGCTGCAAAAATGCCAA 323
Db 324 TGCTGCAAAAATGCCAA 341

RESULT 15
AAC41742/c
ID AAC41742 standard; DNA; 1947 BP.
XX
AC AAC41742;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 32967.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
```

PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 25-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 1.9%; Score 18; DB 21; Length 1947;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 885 TGAACCTTTCACCTCGATGA 902  
| | | | | | | | | | | | | | | |  
Db 1147 TGAACCTTTCACCTCGATGA 1130

Search completed: July 31, 2003, 11:28:37  
Job time : 332 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 11:18:37 ; Search time 61 Seconds  
(without alignments)  
4886.718 Million cell updates/sec

Title: US-09-992-430B-21  
Perfect score: 972  
Sequence: 1 atgttccaagatacaaaagtc.....ttgagcatctggagatcaat 972

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	1.9	1728	4 US-09-277-565-25	Sequence 25, Appl
2	18	1.9	7930	4 US-09-221-017B-356	Sequence 356, App
3	18	1.9	29604	3 US-08-781-891-207	Sequence 207, App
4	17	1.7	193	1 US-08-248-474-44	Sequence 44, Appl
5	17	1.7	193	3 US-08-756-849-44	Sequence 44, Appl
6	17	1.7	402	4 US-09-315-794-61	Sequence 61, Appl
7	17	1.7	402	4 US-09-389-341-61	Sequence 61, Appl
8	17	1.7	965	4 US-09-173-300-16	Sequence 16, Appl
9	17	1.7	993	4 US-09-134-001C-1491	Sequence 1491, Ap
10	17	1.7	1112	1 US-08-136-743B-1	Sequence 1, Appli
11	17	1.7	1240	1 US-08-869-506-1	Sequence 1, Appli
12	17	1.7	1240	3 US-09-128-967-1	Sequence 1, Appli
13	17	1.7	1315	4 US-09-234-186-2	Sequence 2, Appli
14	17	1.7	1315	4 US-09-234-186-4	Sequence 4, Appli
15	17	1.7	1315	4 US-09-234-186-5	Sequence 5, Appli
16	17	1.7	1315	4 US-09-234-186-6	Sequence 6, Appli
17	17	1.7	1315	4 US-09-233-527-2	Sequence 2, Appli
18	17	1.7	1315	4 US-09-233-527-4	Sequence 4, Appli
19	17	1.7	1315	4 US-09-233-527-5	Sequence 5, Appli
20	17	1.7	1315	4 US-09-233-527-6	Sequence 6, Appli
21	17	1.7	1315	5 PCT-US93-05651-2	Sequence 2, Appli
22	17	1.7	1404	3 US-08-961-083-9	Sequence 9, Appli
23	17	1.7	1524	1 US-08-409-122-1	Sequence 1, Appli
24	17	1.7	1524	2 US-08-408-669-1	Sequence 1, Appli
25	17	1.7	1954	4 US-09-370-807-3	Sequence 3, Appli
26	17	1.7	1954	4 US-09-921-259-3	Sequence 3, Appli
27	17	1.7	2387	1 US-08-288-408-4	Sequence 4, Appli

28	17	1.7	2387	4 US-08-655-782-4	Sequence 4, Appli
c 29	17	1.7	4881	4 US-08-961-527-125	Sequence 125, App
30	17	1.7	5600	4 US-08-867-611-3	Sequence 3, Appli
31	17	1.7	5600	5 PCT-US92-06965A-8	Sequence 8, Appli
32	17	1.7	6559	4 US-09-234-186-1	Sequence 1, Appli
33	17	1.7	6559	4 US-09-233-527-1	Sequence 1, Appli
34	17	1.7	6560	5 PCT-US93-05651-1	Sequence 1, Appli
c 35	16	1.6	81	1 US-08-472-255A-166	Sequence 166, App
c 36	16	1.6	81	1 US-08-479-724A-166	Sequence 166, App
c 37	16	1.6	81	3 US-08-472-256B-166	Sequence 166, App
c 38	16	1.6	81	4 US-08-952-793-166	Sequence 166, App
c 39	16	1.6	81	5 PCT-US96-09455A-166	Sequence 166, App
c 40	16	1.6	300	1 US-07-702-163B-5	Sequence 5, Appli
c 41	16	1.6	300	1 US-07-702-163B-6	Sequence 6, Appli
c 42	16	1.6	300	1 US-08-060-984-5	Sequence 5, Appli
c 43	16	1.6	300	1 US-08-060-984-6	Sequence 6, Appli
c 44	16	1.6	300	1 US-08-383-742-5	Sequence 5, Appli
c 45	16	1.6	300	1 US-08-383-742-6	Sequence 6, Appli
c 46	16	1.6	316	2 US-09-032-684-10	Sequence 10, Appl
c 47	16	1.6	650	4 US-09-718-093-5	Sequence 5, Appli
c 48	16	1.6	651	4 US-09-718-093-4	Sequence 4, Appli
c 49	16	1.6	663	1 US-07-624-313-3	Sequence 3, Appli
c 50	16	1.6	960	4 US-08-936-165A-46	Sequence 46, Appl
c 51	16	1.6	1001	4 US-09-641-638-103	Sequence 103, App
c 52	16	1.6	1491	4 US-09-134-001C-866	Sequence 866, App
c 53	16	1.6	1806	2 US-08-819-013-2	Sequence 2, Appli
c 54	16	1.6	2054	4 US-08-981-729-5	Sequence 5, Appli
c 55	16	1.6	2055	1 US-07-842-349-1	Sequence 1, Appli
c 56	16	1.6	2069	1 US-08-619-554-7	Sequence 7, Appli
c 57	16	1.6	2991	3 US-08-795-430-48	Sequence 48, Appl
c 58	16	1.6	2991	4 US-09-355-700-48	Sequence 48, Appl
c 59	16	1.6	3690	3 US-08-991-408-3	Sequence 3, Appli
c 60	16	1.6	3690	4 US-09-432-473-3	Sequence 3, Appli
c 61	16	1.6	5145	3 US-08-991-408-1	Sequence 1, Appli
c 62	16	1.6	5145	4 US-09-432-473-1	Sequence 1, Appli
c 63	16	1.6	6449	2 US-08-479-573A-4	Sequence 4, Appli
c 64	16	1.6	6449	4 US-09-083-116-4	Sequence 4, Appli
c 65	16	1.6	6449	4 US-09-134-916A-4	Sequence 4, Appli
c 66	16	1.6	8473	4 US-09-172-422-2	Sequence 2, Appli
c 67	16	1.6	10236	4 US-09-453-702B-240	Sequence 240, App
c 68	16	1.6	4403765	4 US-09-103-840A-2	Sequence 2, Appli
c 69	16	1.6	4411529	4 US-09-103-840A-1	Sequence 1, Appli
c 70	15	1.5	45	2 US-08-823-516-158	Sequence 158, App
71	15	1.5	282	2 US-09-006-675-5	Sequence 5, Appli
72	15	1.5	282	4 US-09-228-603A-5	Sequence 5, Appli
73	15	1.5	318	4 US-09-134-001C-2112	Sequence 2112, Ap
74	15	1.5	351	4 US-09-602-877A-77	Sequence 77, Appl
c 75	15	1.5	384	4 US-09-004-838-7	Sequence 7, Appli
76	15	1.5	401	4 US-09-221-298-45	Sequence 45, Appl
77	15	1.5	493	4 US-09-222-575-53	Sequence 53, Appl
c 78	15	1.5	535	4 US-09-328-111-298	Sequence 298, App
c 79	15	1.5	696	2 US-08-336-198C-6	Sequence 6, Appli
c 80	15	1.5	789	2 US-08-602-359A-25	Sequence 25, Appl
c 81	15	1.5	822	4 US-09-134-001C-1080	Sequence 1080, Ap
82	15	1.5	837	1 US-08-371-082-1	Sequence 1, Appli
83	15	1.5	855	4 US-09-247-373B-37	Sequence 37, Appl
c 84	15	1.5	946	4 US-09-004-838-4	Sequence 4, Appli
c 85	15	1.5	960	2 US-08-956-047-45	Sequence 45, Appl
86	15	1.5	1001	4 US-09-641-638-383	Sequence 383, App
87	15	1.5	1001	4 US-09-641-638-384	Sequence 384, App
88	15	1.5	1001	4 US-09-641-638-385	Sequence 385, App
89	15	1.5	1001	4 US-09-641-638-386	Sequence 386, App
90	15	1.5	1020	4 US-09-247-155-143	Sequence 143, App
c 91	15	1.5	1036	4 US-09-004-838-3	Sequence 3, Appli
c 92	15	1.5	1074	4 US-09-134-001C-542	Sequence 542, App
c 93	15	1.5	1078	4 US-09-492-543-28	Sequence 28, Appl
c 94	15	1.5	1078	4 US-09-492-543-29	Sequence 29, Appl
c 95	15	1.5	1242	5 PCT-US91-00909-3	Sequence 3, Appli
96	15	1.5	1310	4 US-08-961-527-266	Sequence 266, App
c 97	15	1.5	1311	4 US-09-134-001C-386	Sequence 386, App
98	15	1.5	1383	3 US-08-935-263-3	Sequence 3, Appli
99	15	1.5	1383	4 US-09-594-185-3	Sequence 3, Appli
c 100	15	1.5	1404	4 US-09-134-001C-615	Sequence 615, App



101	15	1.5	1411	4	US-09-387-574-1	Sequence 1, Appli	14	1.4	48	3	US-08-369-822C-37	Sequence 37, Appl
102	15	1.5	1411	4	US-09-668-096-1	Sequence 1, Appli	14	1.4	48	3	US-08-582-776C-52	Sequence 52, Appl
103	15	1.5	1491	2	US-09-006-675-1	Sequence 1, Appli	14	1.4	48	3	US-08-434-831B-49	Sequence 49, Appl
104	15	1.5	1491	4	US-09-228-603A-1	Sequence 1, Appli	14	1.4	50	2	US-08-316-439A-33	Sequence 33, Appl
105	15	1.5	1509	4	US-09-134-001C-761	Sequence 761, App	14	1.4	69	4	US-08-687-421-383	Sequence 383, App
106	15	1.5	1512	3	US-08-476-509B-1	Sequence 1, Appli	14	1.4	70	4	US-08-687-421-378	Sequence 378, App
107	15	1.5	1626	3	US-08-348-518C-1	Sequence 1, Appli	14	1.4	75	1	US-08-447-169A-89	Sequence 89, Appl
108	15	1.5	1762	1	US-07-870-029-1	Sequence 1, Appli	14	1.4	75	2	US-08-233-012C-89	Sequence 89, Appl
109	15	1.5	1882	4	US-09-091-725-14	Sequence 14, Appl	14	1.4	79	1	US-08-471-985A-107	Sequence 107, App
110	15	1.5	1969	1	US-08-106-761-3	Sequence 3, Appli	14	1.4	79	1	US-08-472-255A-163	Sequence 163, App
111	15	1.5	2125	4	US-09-232-160-2	Sequence 2, Appli	14	1.4	79	1	US-08-479-724A-163	Sequence 163, App
112	15	1.5	2187	4	US-09-134-001C-2131	Sequence 2131, Ap	14	1.4	79	3	US-08-472-256B-163	Sequence 163, App
113	15	1.5	2272	1	US-08-233-005-1	Sequence 1, Appli	14	1.4	79	4	US-08-952-793-163	Sequence 163, App
114	15	1.5	2272	1	US-08-428-943-1	Sequence 1, Appli	14	1.4	79	5	PCT-US95-12401A-107	Sequence 107, App
115	15	1.5	2272	3	US-09-016-649-1	Sequence 1, Appli	14	1.4	79	5	PCT-US96-09455A-163	Sequence 163, App
116	15	1.5	2272	5	PCT-US95-04858-1	Sequence 1, Appli	14	1.4	80	1	US-08-471-985A-86	Sequence 86, Appl
117	15	1.5	2357	4	US-09-149-476-262	Sequence 262, App	14	1.4	80	1	US-08-471-985A-88	Sequence 88, Appl
118	15	1.5	2361	4	US-09-011-197-1	Sequence 1, Appli	14	1.4	80	1	US-08-472-255A-132	Sequence 132, App
119	15	1.5	2404	4	US-08-945-771-1	Sequence 1, Appli	14	1.4	80	1	US-08-479-724A-132	Sequence 132, App
120	15	1.5	2514	4	US-09-149-476-261	Sequence 261, App	14	1.4	80	3	US-08-472-256B-132	Sequence 132, App
121	15	1.5	2515	1	US-08-061-465-3	Sequence 3, Appli	14	1.4	80	4	US-08-952-793-132	Sequence 132, App
122	15	1.5	2774	1	US-07-723-002C-5	Sequence 5, Appli	14	1.4	80	5	PCT-US95-12401A-86	Sequence 86, Appl
123	15	1.5	2811	1	US-08-040-548-31	Sequence 31, Appl	14	1.4	80	5	PCT-US95-12401A-88	Sequence 88, Appl
124	15	1.5	2811	1	US-08-466-344-31	Sequence 31, Appl	14	1.4	80	5	PCT-US96-09455A-132	Sequence 132, App
125	15	1.5	2817	6	5206152-6	Patent No. 5206152	14	1.4	81	4	US-08-687-421-379	Sequence 379, App
126	15	1.5	3487	4	US-09-453-702B-164	Sequence 164, App	14	1.4	81	4	US-08-687-421-380	Sequence 380, App
127	15	1.5	3549	4	US-09-381-862-1	Sequence 1, Appli	14	1.4	85	1	US-08-447-169A-201	Sequence 201, App
128	15	1.5	3827	2	US-08-447-031A-1	Sequence 1, Appli	14	1.4	86	1	US-08-447-169A-186	Sequence 186, App
129	15	1.5	3876	5	PCT-US95-17026-1	Sequence 1, Appli	14	1.4	89	4	US-09-133-321-8	Sequence 8, Appli
130	15	1.5	4163	4	US-09-004-838-70	Sequence 70, Appl	14	1.4	90	4	US-08-952-793-176	Sequence 176, App
131	15	1.5	4208	4	US-09-004-838-1	Sequence 1, Appli	14	1.4	90	5	PCT-US96-09455A-176	Sequence 176, App
132	15	1.5	4543	4	US-09-221-017B-1103	Sequence 1103, Ap	14	1.4	145	4	US-09-191-608-24	Sequence 24, Appl
133	15	1.5	4612	2	US-08-447-031A-8	Sequence 8, Appli	14	1.4	236	4	US-09-328-111-27	Sequence 27, Appl
134	15	1.5	4736	4	US-09-056-105-12	Sequence 12, Appl	14	1.4	236	4	US-09-328-111-108	Sequence 108, App
135	15	1.5	4833	4	US-09-066-047-1	Sequence 1, Appli	14	1.4	239	1	US-08-253-155A-50	Sequence 50, Appl
136	15	1.5	5857	4	US-09-293-170-4	Sequence 4, Appli	14	1.4	270	1	US-08-686-878A-31	Sequence 31, Appl
137	15	1.5	5857	4	US-08-961-527-28	Sequence 28, Appl	14	1.4	270	4	US-09-175-928-31	Sequence 31, Appl
138	15	1.5	6038	4	US-09-305-639-4	Sequence 4, Appli	14	1.4	301	4	US-09-605-785-263	Sequence 263, App
139	15	1.5	6519	1	US-08-588-985-1	Sequence 1, Appli	14	1.4	301	4	US-09-605-785-302	Sequence 302, App
140	15	1.5	6519	1	US-08-971-988-1	Sequence 1, Appli	14	1.4	301	4	US-09-439-313-263	Sequence 263, App
141	15	1.5	6822	4	US-09-426-998-3	Sequence 3, Appli	14	1.4	301	4	US-09-439-313-302	Sequence 302, App
142	15	1.5	6828	1	US-08-061-465-1	Sequence 1, Appli	14	1.4	301	4	US-09-352-616A-263	Sequence 263, App
143	15	1.5	7143	4	US-09-381-862-4	Sequence 4, Appli	14	1.4	301	4	US-09-352-616A-302	Sequence 302, App
144	15	1.5	7622	4	US-09-305-639-1	Sequence 1, Appli	14	1.4	301	4	US-09-232-149A-263	Sequence 263, App
145	15	1.5	7741	4	US-09-426-998-4	Sequence 4, Appli	14	1.4	301	4	US-09-232-149A-302	Sequence 302, App
146	15	1.5	8655	3	US-09-075-272-1	Sequence 1, Appli	14	1.4	305	1	US-08-062-472B-5	Sequence 5, Appli
147	15	1.5	9100	2	US-08-743-637B-27	Sequence 27, Appl	14	1.4	308	1	US-08-686-878A-40	Sequence 40, Appl
148	15	1.5	9100	3	US-08-526-840B-27	Sequence 27, Appl	14	1.4	308	4	US-09-175-928-40	Sequence 40, Appl
149	15	1.5	9936	4	US-08-972-927-2	Sequence 2, Appli	14	1.4	312	1	US-08-062-472B-4	Sequence 4, Appli
150	15	1.5	10917	3	US-08-926-842B-11	Sequence 11, Appl	14	1.4	312	4	US-09-134-001C-2436	Sequence 2436, Ap
151	15	1.5	12804	4	US-09-453-702B-149	Sequence 149, App	14	1.4	342	4	US-09-657-453A-18	Sequence 18, Appl
152	15	1.5	13206	4	US-08-961-527-33	Sequence 33, Appl	14	1.4	348	1	US-08-594-031-85	Sequence 85, Appl
153	15	1.5	15363	4	US-08-961-527-139	Sequence 139, App	14	1.4	350	3	US-08-888-077A-35	Sequence 35, Appl
154	15	1.5	17656	4	US-09-433-579-3	Sequence 3, Appli	14	1.4	369	4	US-09-422-487-5	Sequence 5, Appli
155	15	1.5	28001	4	US-09-819-993-3	Sequence 3, Appli	14	1.4	369	4	US-09-378-088A-127	Sequence 127, App
156	15	1.5	34185	4	US-09-545-481-3	Sequence 3, Appli	14	1.4	387	4	US-09-370-838-88	Sequence 88, Appl
157	15	1.5	40000	4	US-09-780-049-18	Sequence 18, Appl	14	1.4	389	4	US-08-943-731-73	Sequence 73, Appl
158	15	1.5	44377	2	US-08-804-227C-7	Sequence 7, Appli	14	1.4	400	4	US-08-976-259-37	Sequence 37, Appl
159	15	1.5	44377	2	US-08-804-198-1	Sequence 1, Appli	14	1.4	406	5	PCT-US91-06234A-4	Sequence 4, Appli
160	15	1.5	50341	1	US-08-247-901C-1	Sequence 1, Appli	14	1.4	417	1	US-08-470-179-117	Sequence 117, App
161	15	1.5	50341	2	US-09-075-904-1	Sequence 1, Appli	14	1.4	420	1	US-08-207-481-30	Sequence 30, Appl
162	15	1.5	52297	4	US-09-426-436-1	Sequence 1, Appli	14	1.4	420	5	PCT-US95-02689-32	Sequence 32, Appl
163	15	1.5	52297	4	US-08-705-557-1	Sequence 1, Appli	14	1.4	437	1	US-08-543-238-1	Sequence 1, Appli
164	15	1.5	65042	4	US-09-784-316-3	Sequence 3, Appli	14	1.4	437	1	US-08-420-526-1	Sequence 1, Appli
165	15	1.5	4403765	4	US-09-103-840A-2	Sequence 2, Appli	14	1.4	445	4	US-09-397-787-149	Sequence 149, App
166	15	1.5	4411529	4	US-09-103-840A-1	Sequence 1, Appli	14	1.4	447	2	US-08-911-445-10	Sequence 10, Appl
167	14	1.4	19	4	US-09-261-115-49	Sequence 49, Appl	14	1.4	447	4	US-09-182-983-10	Sequence 10, Appl
168	14	1.4	20	5	PCT-US95-02080-2	Sequence 2, Appli	14	1.4	471	2	US-08-465-380-11	Sequence 11, Appl
169	14	1.4	27	4	US-09-024-532-8	Sequence 8, Appli	14	1.4	471	2	US-08-480-478-40	Sequence 40, Appl
170	14	1.4	32	1	US-08-279-058B-46	Sequence 46, Appl	14	1.4	471	2	US-08-486-397-11	Sequence 11, Appl
171	14	1.4	32	4	US-08-828-323-46	Sequence 46, Appl	14	1.4	471	2	US-08-486-399-11	Sequence 11, Appl
172	14	1.4	42	1	US-08-270-314-17	Sequence 17, Appl	14	1.4	471	2	US-08-461-965-11	Sequence 11, Appl
173	14	1.4	42	4	US-09-225-302-17	Sequence 17, Appl	14	1.4	471	2	US-08-326-110A-40	Sequence 40, Appl

247	14	1.4	471	2	US-08-634-641-11	Sequence 11, Appl	Sequence 11, Appl	14	320	14	1.4	894	5	PCT-US95-14442A-90	Sequence 90, Appl
248	14	1.4	471	3	US-09-249-471-11	Sequence 11, Appl	Sequence 11, Appl	14	c 321	14	1.4	895	4	US-09-280-116-87	Sequence 87, Appl
249	14	1.4	471	3	US-09-249-472-11	Sequence 11, Appl	Sequence 11, Appl	14	322	14	1.4	910	4	US-09-221-017B-62	Sequence 62, Appl
250	14	1.4	471	3	US-09-249-451-11	Sequence 11, Appl	Sequence 11, Appl	14	c 323	14	1.4	910	4	US-09-221-017B-62	Sequence 62, Appl
251	14	1.4	471	3	US-08-809-455-11	Sequence 11, Appl	Sequence 11, Appl	14	c 324	14	1.4	911	4	US-09-457-046B-11	Sequence 11, Appl
252	14	1.4	471	3	US-09-249-461-11	Sequence 11, Appl	Sequence 11, Appl	14	c 325	14	1.4	917	4	US-09-221-017B-678	Sequence 678, App
253	14	1.4	471	3	US-09-249-448-11	Sequence 11, Appl	Sequence 11, Appl	14	326	14	1.4	944	4	US-09-386-493-4	Sequence 4, Appl
254	14	1.4	472	4	US-09-605-785-351	Sequence 351, App	Sequence 351, App	14	327	14	1.4	973	4	US-08-936-165A-136	Sequence 136, App
255	14	1.4	472	4	US-09-439-313-351	Sequence 351, App	Sequence 351, App	14	328	14	1.4	975	4	US-09-134-001C-528	Sequence 528, App
256	14	1.4	472	4	US-09-352-616A-351	Sequence 351, App	Sequence 351, App	14	329	14	1.4	981	4	US-09-134-001C-437	Sequence 437, App
257	14	1.4	477	3	US-08-867-381A-2	Sequence 2, Appli	Sequence 2, Appli	14	c 330	14	1.4	989	4	US-09-446-504-41	Sequence 41, Appl
258	14	1.4	477	4	US-09-521-144-2	Sequence 2, Appli	Sequence 2, Appli	14	c 331	14	1.4	989	4	US-09-712-266-41	Sequence 41, Appl
259	14	1.4	499	4	US-09-155-036-8	Sequence 8, Appli	Sequence 8, Appli	14	332	14	1.4	1001	4	US-09-641-638-360	Sequence 360, App
260	14	1.4	502	4	US-09-166-350-2	Sequence 2, Appli	Sequence 2, Appli	14	333	14	1.4	1001	4	US-09-641-638-361	Sequence 361, App
c 261	14	1.4	503	4	US-09-004-838-134	Sequence 134, App	Sequence 134, App	14	334	14	1.4	1001	4	US-09-641-638-362	Sequence 362, App
262	14	1.4	503	4	US-09-155-036-11	Sequence 11, Appl	Sequence 11, Appl	14	335	14	1.4	1001	4	US-09-641-638-446	Sequence 446, App
263	14	1.4	507	4	US-09-134-001C-2678	Sequence 2678, Ap	Sequence 2678, Ap	14	c 336	14	1.4	1008	4	US-09-198-955A-9	Sequence 9, Appli
264	14	1.4	515	4	US-09-276-531-99	Sequence 99, Appl	Sequence 99, Appl	14	c 337	14	1.4	1008	4	US-09-694-531-9	Sequence 9, Appli
c 265	14	1.4	518	4	US-08-936-165A-32	Sequence 32, Appl	Sequence 32, Appl	14	338	14	1.4	1029	1	US-08-096-182A-3	Sequence 3, Appli
266	14	1.4	558	4	US-09-134-001C-2099	Sequence 2099, Ap	Sequence 2099, Ap	14	339	14	1.4	1029	1	US-08-877-109-3	Sequence 3, Appli
c 267	14	1.4	580	4	US-09-228-986-23	Sequence 23, Appl	Sequence 23, Appl	14	340	14	1.4	1029	3	US-08-798-760-3	Sequence 3, Appli
c 268	14	1.4	584	4	US-09-328-111-83	Sequence 83, Appl	Sequence 83, Appl	14	341	14	1.4	1029	5	PCT-US94-08327-3	Sequence 3, Appli
c 269	14	1.4	586	4	US-09-385-982-472	Sequence 472, App	Sequence 472, App	14	342	14	1.4	1029	6	5196333-1	Patent No. 5196333
c 270	14	1.4	594	4	US-09-328-111-257	Sequence 257, App	Sequence 257, App	14	343	14	1.4	1032	1	US-08-553-633A-3	Sequence 3, Appli
271	14	1.4	597	4	US-09-385-982-277	Sequence 277, App	Sequence 277, App	14	344	14	1.4	1036	4	US-09-004-838-3	Sequence 3, Appli
c 272	14	1.4	606	4	US-09-032-337-42	Sequence 42, Appl	Sequence 42, Appl	14	345	14	1.4	1038	4	US-09-403-768-3	Sequence 3, Appli
273	14	1.4	607	4	US-09-385-982-525	Sequence 525, App	Sequence 525, App	14	346	14	1.4	1074	4	US-09-347-803-5	Sequence 5, Appli
274	14	1.4	609	4	US-09-385-982-415	Sequence 415, App	Sequence 415, App	14	347	14	1.4	1074	4	US-09-134-001C-382	Sequence 382, App
275	14	1.4	650	4	US-09-328-111-179	Sequence 179, App	Sequence 179, App	14	348	14	1.4	1092	1	US-08-096-182A-5	Sequence 5, Appli
c 276	14	1.4	650	4	US-08-961-527-286	Sequence 286, App	Sequence 286, App	14	349	14	1.4	1092	1	US-08-877-109-5	Sequence 5, Appli
277	14	1.4	654	2	US-08-635-761-89	Sequence 89, Appl	Sequence 89, Appl	14	350	14	1.4	1092	3	US-08-798-760-5	Sequence 5, Appli
278	14	1.4	654	4	US-09-312-520-89	Sequence 89, Appl	Sequence 89, Appl	14	351	14	1.4	1092	5	PCT-US94-08327-5	Sequence 5, Appli
279	14	1.4	664	1	US-08-412-614-89	Sequence 89, Appl	Sequence 89, Appl	14	352	14	1.4	1093	4	US-09-348-265-1	Sequence 5, Appli
c 280	14	1.4	687	4	US-09-040-984-38	Sequence 38, Appl	Sequence 38, Appl	14	c 353	14	1.4	1098	4	US-09-362-473-7	Sequence 7, Appli
c 281	14	1.4	687	4	US-09-643-597-38	Sequence 38, Appl	Sequence 38, Appl	14	354	14	1.4	1100	4	US-07-861-458C-4	Sequence 4, Appli
c 282	14	1.4	687	4	US-09-123-912-38	Sequence 38, Appl	Sequence 38, Appl	14	355	14	1.4	1104	1	US-08-307-499-57	Sequence 57, Appl
283	14	1.4	701	4	US-09-133-321-1	Sequence 1, Appli	Sequence 1, Appli	14	356	14	1.4	1104	1	US-08-423-399B-36	Sequence 36, Appl
284	14	1.4	705	4	US-09-134-001C-1677	Sequence 1677, Ap	Sequence 1677, Ap	14	357	14	1.4	1104	4	US-09-299-268-57	Sequence 57, Appl
c 285	14	1.4	707	1	US-08-062-472B-1	Sequence 1, Appli	Sequence 1, Appli	14	358	14	1.4	1107	2	US-08-933-750C-77	Sequence 77, Appl
286	14	1.4	720	4	US-08-961-527-375	Sequence 375, App	Sequence 375, App	14	359	14	1.4	1107	3	US-09-234-613-77	Sequence 77, Appl
287	14	1.4	729	4	US-09-134-001C-995	Sequence 995, App	Sequence 995, App	14	360	14	1.4	1121	1	US-08-433-783-45	Sequence 45, Appl
c 288	14	1.4	750	4	US-09-446-504-2	Sequence 2, Appli	Sequence 2, Appli	14	361	14	1.4	1121	2	US-08-337-358-45	Sequence 45, Appl
c 289	14	1.4	750	4	US-09-712-266-2	Sequence 2, Appli	Sequence 2, Appli	14	362	14	1.4	1121	5	PCT-US95-07537A-45	Sequence 45, Appl
290	14	1.4	758	4	US-08-943-607-18	Sequence 18, Appl	Sequence 18, Appl	14	363	14	1.4	1121	5	PCT-US95-07537-45	Sequence 45, Appl
291	14	1.4	758	4	US-09-712-016-28	Sequence 28, Appl	Sequence 28, Appl	14	c 364	14	1.4	1134	4	US-09-561-756-13	Sequence 13, Appl
292	14	1.4	759	4	US-08-943-607-20	Sequence 20, Appl	Sequence 20, Appl	14	c 365	14	1.4	1134	4	US-09-227-721-13	Sequence 13, Appl
293	14	1.4	760	4	US-08-943-607-19	Sequence 19, Appl	Sequence 19, Appl	14	366	14	1.4	1160	3	US-08-995-159-1	Sequence 1, Appli
c 294	14	1.4	773	4	US-08-998-416-569	Sequence 569, App	Sequence 569, App	14	367	14	1.4	1162	1	US-08-423-399B-34	Sequence 34, Appl
295	14	1.4	788	4	US-08-991-789A-177	Sequence 177, App	Sequence 177, App	14	368	14	1.4	1162	2	US-08-332-562A-135	Sequence 135, App
296	14	1.4	788	4	US-09-062-451-177	Sequence 177, App	Sequence 177, App	14	c 369	14	1.4	1168	4	US-08-961-527-362	Sequence 362, App
297	14	1.4	788	4	US-09-598-326-177	Sequence 177, App	Sequence 177, App	14	370	14	1.4	1185	4	US-09-134-001C-1656	Sequence 1656, Ap
c 298	14	1.4	794	4	US-08-998-416-455	Sequence 455, App	Sequence 455, App	14	371	14	1.4	1200	3	US-08-867-381A-1	Sequence 1, Appli
c 299	14	1.4	797	1	US-08-173-510B-101	Sequence 101, App	Sequence 101, App	14	372	14	1.4	1200	4	US-09-521-144-1	Sequence 1, Appli
c 300	14	1.4	797	1	US-08-458-218-99	Sequence 99, Appl	Sequence 99, Appl	14	373	14	1.4	1206	4	US-09-191-608-21	Sequence 21, Appl
c 301	14	1.4	797	2	US-08-450-497-101	Sequence 101, App	Sequence 101, App	14	c 374	14	1.4	1211	2	US-08-997-080-40	Sequence 40, Appl
302	14	1.4	799	4	US-09-149-476-288	Sequence 288, App	Sequence 288, App	14	c 375	14	1.4	1211	2	US-08-997-362-40	Sequence 40, Appl
303	14	1.4	807	4	US-09-134-001C-1892	Sequence 1892, Ap	Sequence 1892, Ap	14	c 376	14	1.4	1211	3	US-08-873-970-40	Sequence 40, Appl
c 304	14	1.4	818	4	US-09-366-887A-15	Sequence 15, Appl	Sequence 15, Appl	14	c 377	14	1.4	1211	4	US-09-095-855-40	Sequence 40, Appl
305	14	1.4	823	4	US-08-998-416-551	Sequence 551, App	Sequence 551, App	14	c 378	14	1.4	1211	4	US-08-705-347A-40	Sequence 40, Appl
c 306	14	1.4	870	4	US-08-961-527-256	Sequence 256, App	Sequence 256, App	14	c 379	14	1.4	1211	4	US-09-324-542-40	Sequence 40, Appl
307	14	1.4	878	4	US-08-998-416-291	Sequence 291, App	Sequence 291, App	14	c 380	14	1.4	1211	4	US-09-205-426-40	Sequence 40, Appl
308	14	1.4	885	1	US-08-433-783-40	Sequence 40, Appl	Sequence 40, Appl	14	c 381	14	1.4	1211	4	US-09-200-643-40	Sequence 40, Appl
309	14	1.4	885	2	US-08-337-358-40	Sequence 40, Appl	Sequence 40, Appl	14	c 382	14	1.4	1242	4	US-09-134-001C-544	Sequence 544, App
c 310	14	1.4	885	4	US-09-424-349A-1	Sequence 1, Appli	Sequence 1, Appli	14	c 383	14	1.4	1242	4	US-09-134-001C-2138	Sequence 2138, Ap
311	14	1.4	885	5	PCT-US95-07537A-40	Sequence 40, Appl	Sequence 40, Appl	14	384	14	1.4	1248	3	US-08-910-505-3	Sequence 3, Appli
312	14	1.4	885	5	PCT-US95-07537-40	Sequence 40, Appl	Sequence 40, Appl	14	c 385	14	1.4	1248	4	US-09-105-537-7	Sequence 7, Appli
313	14	1.4	894	3	US-08-906-769-90	Sequence 90, Appl	Sequence 90, Appl	14	386	14	1.4	1251	3	US-08-910-505-1	Sequence 1, Appli
314	14	1.4	894	3	US-08-906-616-90	Sequence 90, Appl	Sequence 90, Appl	14	387	14	1.4	1260	1	US-08-599-252-79	Sequence 79, Appl
315	14	1.4	894	3	US-08-817-795-90	Sequence 90, Appl	Sequence 90, Appl	14	388	14	1.4	1260	1	US-08-436-074-52	Sequence 52, Appl
316	14	1.4	894	3	US-08-639-075A-90	Sequence 90, Appl	Sequence 90, Appl	14	389	14	1.4	1260	5	PCT-US96-06352-79	Sequence 79, Appl
317	14	1.4	894	4	US-09-012-431-90	Sequence 90, Appl	Sequence 90, Appl	14	390	14	1.4	1260	5	PCT-US96-06583-79	Sequence 79, Appl
318	14	1.4	894	4	US-09-012-692-90	Sequence 90, Appl	Sequence 90, Appl	14	c 391	14	1.4	1266	4	US-09-199-637A-115	Sequence 115, App
319	14	1.4	894	4	US-08-906-613-90	Sequence 90, Appl	Sequence 90, Appl	14	c 392	14	1.4	1269	5	PCT-US93-00893-4	Sequence 4, Appli

393	14	1.4	1272	2	US-08-972-258-3	Sequence 3, Appli	466	14	1.4	1621	4	US-09-292-225-36	Sequence 36, Appli
394	14	1.4	1272	4	US-09-263-128-3	Sequence 3, Appli	467	14	1.4	1624	1	US-08-416-870C-5	Sequence 5, Appli
395	14	1.4	1275	2	US-08-911-445-4	Sequence 4, Appli	468	14	1.4	1633	4	US-09-155-036-17	Sequence 17, Appl
396	14	1.4	1275	4	US-09-182-983-4	Sequence 1, Appli	469	14	1.4	1652	4	US-08-858-207A-62	Sequence 62, Appl
C 397	14	1.4	1291	3	US-08-776-980C-1	Sequence 1, Appli	C 470	14	1.4	1666	4	US-09-360-197-5	Sequence 5, Appli
C 398	14	1.4	1291	4	US-09-268-195C-1	Sequence 1, Appli	471	14	1.4	1674	4	US-09-155-036-21	Sequence 21, Appl
C 399	14	1.4	1297	1	US-08-370-975B-12	Sequence 12, Appl	472	14	1.4	1675	4	US-09-329-535-1	Sequence 1, Appli
C 400	14	1.4	1299	3	US-08-721-986-8	Sequence 8, Appli	473	14	1.4	1692	1	US-08-451-715A-9	Sequence 9, Appli
C 401	14	1.4	1299	3	US-08-225-487A-8	Sequence 8, Appli	C 474	14	1.4	1699	2	US-08-484-993B-3	Sequence 3, Appli
C 402	14	1.4	1299	4	US-09-199-637A-160	Sequence 160, App	C 475	14	1.4	1699	2	US-08-484-158B-3	Sequence 3, Appli
C 403	14	1.4	1318	5	PCT-US94-07127A-1	Sequence 1, Appli	C 476	14	1.4	1699	2	US-08-484-596A-3	Sequence 3, Appli
C 404	14	1.4	1320	4	US-09-457-046B-44	Sequence 44, Appl	C 477	14	1.4	1699	2	US-08-480-150A-3	Sequence 3, Appli
C 405	14	1.4	1329	4	US-09-305-984-13	Sequence 13, Appl	C 478	14	1.4	1699	3	US-08-458-731-3	Sequence 3, Appli
C 406	14	1.4	1329	4	US-09-073-541A-13	Sequence 13, Appl	C 479	14	1.4	1699	3	US-08-149-223A-3	Sequence 3, Appli
407	14	1.4	1353	1	US-08-625-322-3	Sequence 3, Appli	C 480	14	1.4	1704	4	US-09-032-337-38	Sequence 38, Appl
C 408	14	1.4	1359	1	US-07-618-312A-1	Sequence 1, Appli	481	14	1.4	1715	1	US-07-847-743B-24	Sequence 24, Appl
C 409	14	1.4	1359	1	US-07-618-312A-3	Sequence 3, Appli	482	14	1.4	1715	1	US-08-456-201-24	Sequence 24, Appl
C 410	14	1.4	1359	1	US-08-110-786A-7	Sequence 7, Appli	483	14	1.4	1715	2	US-08-456-241-24	Sequence 24, Appl
C 411	14	1.4	1359	1	US-08-280-228-1	Sequence 1, Appli	484	14	1.4	1715	5	PCT-US92-04295A-24	Sequence 24, Appl
C 412	14	1.4	1359	1	US-08-280-228-3	Sequence 3, Appli	485	14	1.4	1722	1	US-08-261-822A-9	Sequence 9, Appli
C 413	14	1.4	1359	4	US-09-134-001C-1336	Sequence 1336, Ap	486	14	1.4	1722	5	PCT-US95-07744A-9	Sequence 9, Appli
414	14	1.4	1380	1	US-08-484-324-1	Sequence 1, Appli	C 487	14	1.4	1734	4	US-09-484-970B-63	Sequence 63, Appl
415	14	1.4	1380	1	US-08-486-409-1	Sequence 1, Appli	488	14	1.4	1753	6	5225348-2	Patent No. 5225348
416	14	1.4	1380	2	US-08-825-784-1	Sequence 1, Appli	489	14	1.4	1762	2	US-08-742-621-2	Sequence 2, Appli
417	14	1.4	1380	5	PCT-US92-01899-1	Sequence 1, Appli	C 490	14	1.4	1764	4	US-09-351-224E-6	Sequence 6, Appli
C 418	14	1.4	1387	3	US-08-868-594-1	Sequence 1, Appli	491	14	1.4	1796	2	US-08-911-445-17	Sequence 17, Appl
C 419	14	1.4	1387	4	US-09-434-323-1	Sequence 1, Appli	492	14	1.4	1796	4	US-09-182-983-17	Sequence 17, Appl
C 420	14	1.4	1414	1	US-08-024-868-1	Sequence 1, Appli	C 493	14	1.4	1800	3	US-08-617-860B-23	Sequence 23, Appl
C 421	14	1.4	1414	2	US-08-242-097-1	Sequence 1, Appli	C 494	14	1.4	1800	4	US-09-643-597-345	Sequence 345, App
C 422	14	1.4	1414	4	US-09-206-695-1	Sequence 1, Appli	495	14	1.4	1808	4	US-09-155-036-3	Sequence 3, Appli
C 423	14	1.4	1430	1	US-08-276-452A-25	Sequence 25, Appl	496	14	1.4	1845	2	US-08-911-445-5	Sequence 5, Appli
C 424	14	1.4	1430	2	US-08-798-744-25	Sequence 25, Appl	497	14	1.4	1845	4	US-09-182-983-5	Sequence 5, Appli
425	14	1.4	1452	1	US-08-149-093A-3	Sequence 3, Appli	498	14	1.4	1846	4	US-09-155-036-4	Sequence 4, Appli
426	14	1.4	1452	1	US-08-911-245-3	Sequence 3, Appli	499	14	1.4	1847	4	US-09-149-476-149	Sequence 149, App
427	14	1.4	1452	1	US-08-553-058C-3	Sequence 3, Appli	C 500	14	1.4	1858	1	US-08-668-381A-6	Sequence 6, Appli
428	14	1.4	1452	2	US-08-514-451A-3	Sequence 3, Appli	501	14	1.4	1858	2	US-08-909-965C-11	Sequence 11, Appl
429	14	1.4	1452	3	US-09-170-331-3	Sequence 3, Appli	502	14	1.4	1858	4	US-09-155-036-1	Sequence 1, Appli
430	14	1.4	1452	4	US-09-510-473-3	Sequence 3, Appli	503	14	1.4	1858	4	US-08-546-793A-1	Sequence 1, Appli
431	14	1.4	1452	4	US-09-048-916B-3	Sequence 3, Appli	C 504	14	1.4	1897	6	RE34606-5	Patent No. RE34,60
432	14	1.4	1455	1	US-08-423-399B-32	Sequence 32, Appl	C 505	14	1.4	1910	1	US-08-247-902A-1	Sequence 1, Appli
433	14	1.4	1458	2	US-08-911-445-19	Sequence 19, Appl	C 506	14	1.4	1910	5	PCT-US93-10541-1	Sequence 1, Appli
434	14	1.4	1458	4	US-09-182-983-19	Sequence 19, Appl	507	14	1.4	1915	1	US-08-496-855A-5	Sequence 5, Appli
C 435	14	1.4	1470	4	US-09-292-225-40	Sequence 40, Appl	508	14	1.4	1915	2	US-08-466-589-11	Sequence 11, Appl
436	14	1.4	1470	4	US-09-292-225-42	Sequence 42, Appl	509	14	1.4	1915	2	US-08-700-636-11	Sequence 11, Appl
437	14	1.4	1481	1	US-08-136-922-1	Sequence 1, Appli	510	14	1.4	1915	3	US-08-467-574-11	Sequence 11, Appl
438	14	1.4	1485	1	US-08-471-601-23	Sequence 23, Appl	511	14	1.4	1915	4	US-09-217-345-11	Sequence 11, Appl
439	14	1.4	1485	1	US-08-474-556-23	Sequence 23, Appl	512	14	1.4	1915	4	US-08-487-596-17	Sequence 17, Appl
440	14	1.4	1485	1	US-08-479-382-23	Sequence 23, Appl	C 513	14	1.4	1927	4	US-08-952-365-3	Sequence 3, Appli
441	14	1.4	1485	1	US-08-470-354-23	Sequence 23, Appl	514	14	1.4	1958	4	US-09-215-221-9	Sequence 9, Appli
442	14	1.4	1485	1	US-08-479-383-23	Sequence 23, Appl	515	14	1.4	1978	4	US-09-198-603C-5	Sequence 5, Appli
443	14	1.4	1485	2	US-08-479-041-23	Sequence 23, Appl	516	14	1.4	1997	2	US-08-750-134A-6	Sequence 6, Appli
444	14	1.4	1485	4	US-08-819-646-23	Sequence 23, Appl	517	14	1.4	1997	4	US-09-363-745-6	Sequence 6, Appli
445	14	1.4	1485	4	US-09-195-716-23	Sequence 23, Appl	C 518	14	1.4	2002	6	5256558-3	Patent No. 5256558
C 446	14	1.4	1513	2	US-08-757-653-159	Sequence 159, App	519	14	1.4	2012	2	US-08-555-568B-16	Sequence 16, Appl
C 447	14	1.4	1513	4	US-08-520-946-159	Sequence 159, App	520	14	1.4	2012	4	US-09-519-223-16	Sequence 16, Appl
C 448	14	1.4	1526	4	US-09-081-689-1	Sequence 1, Appli	521	14	1.4	2023	4	US-09-288-143-54	Sequence 54, Appl
C 449	14	1.4	1527	4	US-09-292-225-37	Sequence 37, Appl	522	14	1.4	2042	2	US-08-933-821-16	Sequence 16, Appl
450	14	1.4	1527	4	US-09-292-225-39	Sequence 39, Appl	523	14	1.4	2042	3	US-08-934-494-5	Sequence 5, Appli
C 451	14	1.4	1535	3	US-09-032-372-8	Sequence 8, Appli	524	14	1.4	2042	3	US-08-960-507-16	Sequence 16, Appl
C 452	14	1.4	1539	2	US-08-828-596-1	Sequence 1, Appli	525	14	1.4	2042	3	US-09-143-068-5	Sequence 5, Appli
C 453	14	1.4	1547	4	US-09-356-818A-1	Sequence 1, Appli	526	14	1.4	2042	4	US-09-143-707-5	Sequence 5, Appli
C 454	14	1.4	1555	4	US-09-227-357-25	Sequence 25, Appl	527	14	1.4	2042	4	US-09-202-089-5	Sequence 5, Appli
455	14	1.4	1567	3	US-08-889-108-16	Sequence 16, Appl	528	14	1.4	2042	4	US-09-136-828-16	Sequence 16, Appl
456	14	1.4	1567	5	PCT-US94-10358-16	Sequence 16, Appl	529	14	1.4	2042	4	US-09-332-928A-16	Sequence 16, Appl
457	14	1.4	1569	3	US-08-821-984-9	Sequence 9, Appli	530	14	1.4	2042	4	US-09-511-133-5	Sequence 5, Appli
458	14	1.4	1569	4	US-09-329-749-9	Sequence 9, Appli	531	14	1.4	2042	4	US-09-136-801-16	Sequence 16, Appl
C 459	14	1.4	1578	4	US-09-351-224E-7	Sequence 7, Appli	532	14	1.4	2042	4	US-09-332-929-16	Sequence 16, Appl
460	14	1.4	1586	4	US-09-155-036-19	Sequence 19, Appl	533	14	1.4	2042	4	US-09-690-169-5	Sequence 5, Appli
461	14	1.4	1588	4	US-08-952-089A-26	Sequence 26, Appl	534	14	1.4	2042	4	US-09-511-631-5	Sequence 5, Appli
462	14	1.4	1597	4	US-09-155-036-20	Sequence 20, Appl	535	14	1.4	2046	2	US-08-911-445-6	Sequence 6, Appli
463	14	1.4	1598	4	US-09-155-036-18	Sequence 18, Appl	536	14	1.4	2046	3	US-09-188-930-214	Sequence 214, App
C 464	14	1.4	1607	2	US-08-797-226-1	Sequence 1, Appli	537	14	1.4	2046	4	US-09-182-983-6	Sequence 6, Appli
C 465	14	1.4	1621	4	US-09-292-225-34	Sequence 34, Appl	538	14	1.4	2048	1	US-08-416-870C-7	Sequence 7, Appli







685	1.4	4084	2	US-08-487-826B-1	Sequence 1, Appli	758	14	1.4	5970	3	US-09-320-878-21	Sequence 21, Appli
686	1.4	4084	4	US-09-210-288-1	Sequence 1, Appli	c 759	14	1.4	5977	3	US-09-024-020B-1	Sequence 1, Appli
687	1.4	4084	6	5198347-5	Patent No. 5198347	c 760	14	1.4	5977	4	US-09-425-043-1	Sequence 1, Appli
688	1.4	4090	4	US-08-569-214-4	Sequence 4, Appli	c 761	14	1.4	6007	3	US-09-024-020B-2	Sequence 2, Appli
c 689	1.4	4090	4	US-08-569-214-4	Sequence 4, Appli	c 762	14	1.4	6007	4	US-09-425-043-2	Sequence 2, Appli
690	1.4	4090	4	US-08-937-236-4	Sequence 4, Appli	c 763	14	1.4	6091	4	US-09-453-702B-200	Sequence 200, App
c 691	1.4	4090	4	US-08-937-236-4	Sequence 4, Appli	764	14	1.4	6111	4	US-09-538-414-9	Sequence 9, Appli
c 692	1.4	4092	4	US-09-306-595C-5	Sequence 5, Appli	c 765	14	1.4	6156	4	US-08-891-640-1	Sequence 1, Appli
693	1.4	4146	4	US-08-952-089A-29	Sequence 29, Appl	c 766	14	1.4	6156	4	US-09-723-535-3	Sequence 3, Appli
c 694	1.4	4185	4	US-08-961-527-173	Sequence 173, App	c 767	14	1.4	6180	1	US-08-386-727-1	Sequence 1, Appli
695	1.4	4211	4	US-09-004-838-106	Sequence 106, App	c 768	14	1.4	6180	2	US-08-600-452A-1	Sequence 1, Appli
696	1.4	4281	4	US-09-357-206A-8	Sequence 8, Appli	c 769	14	1.4	6202	1	US-08-484-101B-41	Sequence 41, Appl
c 697	1.4	4392	4	US-09-657-453A-17	Sequence 17, Appl	c 770	14	1.4	6202	4	US-08-714-524D-41	Sequence 41, Appl
698	1.4	4454	2	US-08-738-172-2	Sequence 2, Appli	c 771	14	1.4	6328	4	US-08-913-832A-1	Sequence 1, Appli
699	1.4	4500	1	US-08-468-557-3	Sequence 3, Appli	c 772	14	1.4	6328	4	US-08-913-832A-1	Sequence 1, Appli
c 700	1.4	4508	5	PCT-US93-06251-34	Sequence 34, Appl	773	14	1.4	6328	4	US-09-249-181A-1	Sequence 1, Appli
701	1.4	4512	1	US-08-224-391-52	Sequence 52, Appl	c 774	14	1.4	6328	4	US-09-249-181A-1	Sequence 1, Appli
702	1.4	4512	1	US-08-484-304-52	Sequence 52, Appl	c 775	14	1.4	6407	2	US-08-616-844-7	Sequence 7, Appli
c 703	1.4	4512	4	US-09-350-982C-3	Sequence 3, Appli	c 776	14	1.4	6407	2	US-08-599-654-7	Sequence 7, Appli
c 704	1.4	4565	3	US-08-776-265-1	Sequence 1, Appli	c 777	14	1.4	6407	3	US-08-944-868A-7	Sequence 7, Appli
c 705	1.4	4566	1	US-08-484-101B-37	Sequence 37, Appl	c 778	14	1.4	6407	3	US-08-944-423A-7	Sequence 7, Appli
c 706	1.4	4566	4	US-09-657-453A-3	Sequence 3, Appli	c 779	14	1.4	6407	3	US-08-944-496-7	Sequence 7, Appli
c 707	1.4	4567	4	US-08-714-524D-37	Sequence 37, Appl	780	14	1.4	6471	4	US-09-353-585-1	Sequence 1, Appli
708	1.4	4588	4	US-09-453-702B-26	Sequence 26, Appl	781	14	1.4	6513	1	US-08-338-702-7	Sequence 7, Appli
709	1.4	4695	4	US-09-309-572-9	Sequence 9, Appli	782	14	1.4	6513	1	US-08-337-339-7	Sequence 7, Appli
710	1.4	4695	6	5225348-3	Patent No. 5225348	783	14	1.4	6513	1	US-08-724-095-7	Sequence 7, Appli
c 711	1.4	4741	4	US-09-056-105-11	Sequence 11, Appl	784	14	1.4	6513	5	PCT-US95-14262-7	Sequence 7, Appli
712	1.4	4755	4	US-09-837-863-24	Sequence 24, Appl	785	14	1.4	6513	5	PCT-US95-14378-7	Sequence 7, Appli
c 713	1.4	4781	2	US-09-001-273-1	Sequence 1, Appli	786	14	1.4	6519	2	US-08-808-793-24	Sequence 24, Appl
c 714	1.4	4781	4	US-08-843-459A-1	Sequence 1, Appli	c 787	14	1.4	6556	3	US-09-024-020B-7	Sequence 7, Appli
715	1.4	4837	1	US-08-089-986-1	Sequence 1, Appli	c 788	14	1.4	6556	4	US-09-425-043-7	Sequence 7, Appli
716	1.4	4837	1	US-08-478-585-1	Sequence 1, Appli	c 789	14	1.4	6586	3	US-09-024-020B-43	Sequence 43, Appl
717	1.4	4837	1	US-08-717-312-1	Sequence 1, Appli	c 790	14	1.4	6586	4	US-09-425-043-43	Sequence 43, Appl
718	1.4	4837	2	US-08-266-408-1	Sequence 1, Appli	c 791	14	1.4	6677	4	US-08-939-366-27	Sequence 27, Appl
719	1.4	4837	5	PCT-US94-07886-1	Sequence 1, Appli	c 792	14	1.4	6677	4	US-09-467-997-6	Sequence 6, Appli
c 720	1.4	4847	3	US-09-061-400-1	Sequence 1, Appli	793	14	1.4	6755	3	US-08-931-999-4	Sequence 4, Appli
721	1.4	4887	1	US-07-789-915A-7	Sequence 7, Appli	794	14	1.4	6800	2	US-08-928-692-25	Sequence 25, Appl
722	1.4	4887	1	US-08-005-002C-7	Sequence 7, Appli	795	14	1.4	6800	4	US-09-339-972-25	Sequence 25, Appl
723	1.4	4887	1	US-08-487-203A-7	Sequence 7, Appli	c 796	14	1.4	6826	3	US-09-024-020B-8	Sequence 8, Appli
c 724	1.4	4899	2	US-08-210-535-5	Sequence 5, Appli	c 797	14	1.4	6826	4	US-09-425-043-8	Sequence 8, Appli
c 725	1.4	4899	6	5268270-1	Patent No. 5268270	798	14	1.4	6854	1	US-08-468-036-4	Sequence 4, Appli
726	1.4	4919	4	US-08-456-200B-2	Sequence 2, Appli	799	14	1.4	6854	2	US-08-376-843-4	Sequence 4, Appli
727	1.4	4977	4	US-09-071-035-257	Sequence 257, App	800	14	1.4	6987	4	US-09-026-033-3	Sequence 3, Appli
728	1.4	4977	4	US-09-071-035-261	Sequence 261, App	c 801	14	1.4	6990	4	US-09-026-033-23	Sequence 23, Appl
729	1.4	4977	4	US-09-071-035-265	Sequence 265, App	802	14	1.4	7032	2	US-08-149-097D-24	Sequence 24, Appl
c 730	1.4	4989	4	US-09-560-639-3	Sequence 3, Appli	803	14	1.4	7032	3	US-08-949-386-24	Sequence 24, Appl
731	1.4	5000	3	US-09-104-070-3	Sequence 3, Appli	804	14	1.4	7032	3	US-08-450-562-24	Sequence 24, Appl
732	1.4	5183	1	US-08-459-568-3	Sequence 3, Appli	805	14	1.4	7032	4	US-08-984-709A-24	Sequence 24, Appl
733	1.4	5183	2	US-08-399-411-3	Sequence 3, Appli	806	14	1.4	7032	4	US-08-450-272-24	Sequence 24, Appl
c 734	1.4	5197	1	US-08-131-365B-53	Sequence 53, Appl	807	14	1.4	7082	4	US-09-362-831-1	Sequence 1, Appli
c 735	1.4	5197	2	US-08-668-123-53	Sequence 53, Appl	808	14	1.4	7089	3	US-08-949-386-25	Sequence 25, Appl
736	1.4	5295	4	US-09-221-017B-27	Sequence 27, Appl	809	14	1.4	7089	3	US-08-450-562-25	Sequence 25, Appl
737	1.4	5434	2	US-08-841-349-1	Sequence 1, Appli	c 810	14	1.4	7089	4	US-08-984-709A-25	Sequence 25, Appl
738	1.4	5461	3	US-07-998-289B-7	Sequence 7, Appli	811	14	1.4	7089	4	US-08-450-272-25	Sequence 25, Appl
c 739	1.4	5521	4	US-08-975-762-48	Sequence 48, Appl	c 812	14	1.4	7091	4	US-08-975-762-46	Sequence 46, Appl
c 740	1.4	5521	4	US-09-295-028-48	Sequence 48, Appl	c 813	14	1.4	7091	4	US-09-295-028-46	Sequence 46, Appl
c 741	1.4	5521	4	US-09-106-582-48	Sequence 48, Appl	c 814	14	1.4	7091	4	US-09-106-582-46	Sequence 46, Appl
742	1.4	5565	3	US-09-068-195-21	Sequence 21, Appl	c 815	14	1.4	7096	4	US-09-221-017B-373	Sequence 373, App
c 743	1.4	5599	2	US-08-477-451-9	Sequence 9, Appli	c 816	14	1.4	7571	4	US-08-961-527-2	Sequence 2, Appli
744	1.4	5599	2	US-08-477-451-13	Sequence 13, Appl	817	14	1.4	7577	4	US-08-961-527-46	Sequence 46, Appl
745	1.4	5605	4	US-09-268-140-6	Sequence 6, Appli	818	14	1.4	8157	1	US-08-358-160-65	Sequence 65, Appl
746	1.4	5857	4	US-08-961-527-28	Sequence 28, Appl	819	14	1.4	8341	2	US-08-545-860D-63	Sequence 63, Appl
747	1.4	5868	3	US-08-516-859A-3	Sequence 3, Appli	820	14	1.4	8342	3	US-08-545-860D-63	Sequence 63, Appl
748	1.4	5868	4	US-09-586-472-3	Sequence 3, Appli	821	14	1.4	8342	5	PCT-US94-04496-63	Sequence 63, Appl
749	1.4	5868	4	US-09-528-706-3	Sequence 3, Appli	c 822	14	1.4	8378	5	PCT-US91-09055-1	Sequence 1, Appli
750	1.4	5889	1	US-08-463-092B-5	Sequence 5, Appli	823	14	1.4	8392	1	US-08-080-255-6	Sequence 6, Appli
751	1.4	5889	2	US-08-462-109A-5	Sequence 5, Appli	824	14	1.4	8392	3	US-08-465-713-6	Sequence 6, Appli
752	1.4	5889	2	US-08-460-907B-5	Sequence 5, Appli	825	14	1.4	8392	5	PCT-US93-05857-6	Sequence 6, Appli
753	1.4	5889	3	US-08-463-179A-5	Sequence 5, Appli	c 826	14	1.4	8442	4	US-09-272-032-6	Sequence 6, Appli
754	1.4	5889	3	US-08-461-384B-5	Sequence 5, Appli	827	14	1.4	8584	1	US-08-358-160-66	Sequence 66, Appl
c 755	1.4	5925	3	US-08-470-260-4	Sequence 4, Appli	828	14	1.4	8590	1	US-08-358-160-70	Sequence 70, Appl
c 756	1.4	5925	3	US-08-471-491-4	Sequence 4, Appli	c 829	14	1.4	8600	4	US-09-457-708-1	Sequence 1, Appli
c 757	1.4	5925	3	US-08-466-662-4	Sequence 4, Appli	830	14	1.4	8697	4	US-08-961-527-123	Sequence 123, App

831 1.4 8789 1 US-08-328-254-5 Sequence 5, Appli  
C 832 1.4 8900 4 US-09-305-984-25 Sequence 25, Appl  
833 1.4 8900 4 US-09-073-541A-25 Sequence 25, Appl  
834 1.4 9058 4 US-08-913-014A-9 Sequence 9, Appli  
C 835 1.4 9461 4 US-09-221-017B-513 Sequence 513, App  
C 836 1.4 9562 3 US-08-886-967-1 Sequence 1, Appli  
C 837 1.4 9562 4 US-09-306-949-1 Sequence 1, Appli  
C 838 1.4 9707 4 US-08-961-527-164 Sequence 164, App  
839 1.4 10136 1 US-08-353-700-2 Sequence 2, Appli  
840 1.4 10136 5 PCT-US95-16216-2 Sequence 2, Appli  
C 841 1.4 10299 2 US-08-477-451-1 Sequence 1, Appli  
842 1.4 10299 2 US-08-477-451-5 Sequence 5, Appli  
843 1.4 10718 3 US-08-325-426B-1 Sequence 1, Appli  
844 1.4 10881 4 US-09-357-206A-9 Sequence 9, Appli  
C 845 1.4 11056 4 US-09-004-838-23 Sequence 23, Appl  
C 846 1.4 11303 4 US-08-961-527-115 Sequence 115, App  
C 847 1.4 11459 4 US-09-454-721A-3 Sequence 3, Appli  
848 1.4 11492 4 US-08-991-840A-1 Sequence 1, Appli  
C 849 1.4 11613 4 US-09-453-702B-42 Sequence 42, Appl  
850 1.4 12286 4 US-09-357-206A-1 Patent No. 5206163  
851 1.4 12492 6 5206163-2 Sequence 91, Appl  
C 852 1.4 12730 4 US-09-004-838-91 Sequence 1, Appli  
853 1.4 12734 4 US-09-344-456-1 Sequence 1, Appli  
C 854 1.4 12793 4 US-09-004-838-124 Sequence 124, App  
855 1.4 12949 4 US-09-538-414-11 Sequence 11, Appl  
856 1.4 13613 4 US-09-105-537-3 Sequence 3, Appli  
857 1.4 14078 4 US-09-433-262-1 Sequence 1, Appli  
858 1.4 14078 4 US-09-702-330-1 Sequence 1, Appli  
859 1.4 14507 3 US-08-785-150-1 Sequence 1, Appli  
860 1.4 14507 4 US-09-660-299-1 Sequence 1, Appli  
861 1.4 14507 4 US-09-435-377-1 Sequence 1, Appli  
862 1.4 14578 3 US-08-859-694-1 Sequence 1, Appli  
C 863 1.4 15062 4 US-09-004-838-89 Sequence 89, Appl  
864 1.4 15213 4 US-08-961-527-26 Sequence 26, Appl  
865 1.4 15222 2 US-08-801-898A-23 Sequence 12, Appl  
C 866 1.4 15222 4 US-08-962-690-12 Sequence 12, Appl  
867 1.4 15223 2 US-08-892-403A-1 Sequence 1, Appli  
868 1.4 15223 4 US-08-720-132-1 Sequence 1, Appli  
869 1.4 15225 2 US-08-892-403A-2 Sequence 2, Appli  
C 870 1.4 19307 3 US-08-836-022A-10 Sequence 10, Appl  
C 871 1.4 19307 4 US-09-427-048A-10 Sequence 10, Appl  
C 872 1.4 19446 4 US-08-961-527-51 Sequence 51, Appl  
C 873 1.4 19702 4 US-08-961-527-7 Sequence 7, Appli  
C 874 1.4 19932 2 US-08-477-451-25 Sequence 25, Appl  
875 1.4 20199 4 US-08-961-527-6 Sequence 6, Appli  
C 876 1.4 25165 4 US-09-453-702B-39 Sequence 39, Appl  
C 877 1.4 26764 1 US-08-370-975B-1 Sequence 1, Appli  
C 878 1.4 28882 4 US-08-961-527-140 Sequence 140, App  
C 879 1.4 31328 4 US-09-215-694-19 Sequence 19, Appl  
880 1.4 31960 4 US-09-453-702B-11 Sequence 11, Appl  
C 881 1.4 36159 4 US-09-749-588-3 Sequence 3, Appli  
882 1.4 37895 1 US-08-375-709-1 Sequence 1, Appli  
883 1.4 37895 1 US-08-752-929-1 Sequence 1, Appli  
884 1.4 37895 3 US-09-090-793-1 Sequence 1, Appli  
885 1.4 37948 4 US-09-251-645-11 Sequence 11, Appl  
C 886 1.4 38584 4 US-09-453-702B-50 Sequence 50, Appl  
887 1.4 38682 4 US-08-943-731-2 Sequence 2, Appli  
C 888 1.4 43950 4 US-09-735-934A-3 Sequence 3, Appli  
C 889 1.4 45716 4 US-08-965-048-5 Sequence 5, Appli  
C 890 1.4 45989 4 US-08-965-048-6 Sequence 6, Appli  
891 1.4 46899 1 US-08-471-119A-1 Sequence 1, Appli  
C 892 1.4 48908 4 US-09-453-702B-137 Sequence 137, App  
893 1.4 48974 4 US-08-920-422-17 Sequence 17, Appl  
894 1.4 50000 4 US-09-146-053-4 Sequence 4, Appli  
C 895 1.4 51952 3 US-08-947-823-1 Sequence 1, Appli  
C 896 1.4 59065 4 US-09-813-817-3 Sequence 3, Appli  
C 897 1.4 59065 4 US-09-978-197-3 Sequence 3, Appli  
898 1.4 61663 4 US-09-453-702B-62 Sequence 62, Appl  
899 1.4 65042 4 US-09-784-316-3 Sequence 3, Appli  
900 1.4 68750 3 US-09-335-409-1 Sequence 1, Appli  
901 1.4 68750 4 US-09-568-102-1 Sequence 1, Appli  
902 1.4 68750 4 US-09-567-969-1 Sequence 1, Appli  
903 1.4 68750 4 US-09-568-480-1 Sequence 1, Appli

14 904 14 68750 4 US-09-568-486-1 Sequence 1, Appli  
14 905 14 68750 4 US-09-568-472-1 Sequence 1, Appli  
14 906 14 68750 4 US-09-567-899-1 Sequence 1, Appli  
14 907 14 70000 4 US-09-851-896-3 Sequence 3, Appli  
14 908 14 71989 4 US-09-443-501A-2 Sequence 2, Appli  
C 909 14 72928 3 US-09-009-913-1 Sequence 1, Appli  
C 910 14 80246 4 US-09-078-294-4 Sequence 4, Appli  
C 911 14 80595 4 US-09-078-294-3 Sequence 3, Appli  
912 14 81001 4 US-09-750-580-1 Sequence 1, Appli  
913 14 84495 4 US-09-797-906-3 Sequence 3, Appli  
914 14 87350 3 US-08-781-891-79 Sequence 79, Appl  
C 915 14 87350 3 US-08-781-891-79 Sequence 79, Appl  
916 14 87543 4 US-09-791-211-3 Sequence 3, Appli  
C 917 14 87543 4 US-09-791-211-3 Sequence 3, Appli  
918 14 98844 4 US-09-791-211-10 Sequence 10, Appl  
C 919 14 111282 4 US-09-754-250-3 Sequence 3, Appli  
C 920 14 162450 4 US-09-345-882-1 Sequence 1, Appli  
921 14 168575 4 US-09-426-290-1 Sequence 1, Appli  
C 922 13 18 4 US-09-193-377B-57 Sequence 57, Appl  
C 923 13 20 4 US-09-710-200-28 Sequence 28, Appl  
C 924 13 21 3 US-08-996-306-63 Sequence 63, Appl  
C 925 13 21 3 US-08-513-974B-256 Sequence 256, App  
926 13 21 3 US-08-823-120-12 Sequence 12, Appl  
927 13 21 4 US-09-338-907-63 Sequence 63, Appl  
928 13 21 4 US-09-218-207-63 Sequence 63, Appl  
C 929 13 24 3 US-08-513-974B-237 Sequence 237, App  
C 930 13 24 4 US-09-439-000-1 Sequence 1, Appli  
C 931 13 25 4 US-09-050-159-106 Sequence 106, App  
C 932 13 26 2 US-08-602-093-7 Sequence 7, Appli  
933 13 27 1 US-08-105-483-366 Sequence 366, App  
C 934 13 27 1 US-08-150-331-33 Sequence 33, Appl  
C 935 13 27 1 US-08-447-313-2 Sequence 2, Appli  
936 13 27 1 US-08-709-209-366 Sequence 366, App  
937 13 27 1 US-08-458-101-366 Sequence 366, App  
C 938 13 27 2 US-08-887-434-2 Sequence 2, Appli  
C 939 13 27 2 US-08-898-560-10 Sequence 10, Appl  
C 940 13 27 2 US-08-898-560-11 Sequence 11, Appl  
C 941 13 27 4 US-08-584-040-783 Sequence 783, App  
C 942 13 30 1 US-08-766-014-14 Sequence 14, Appl  
C 943 13 31 4 US-09-183-861-28 Sequence 28, Appl  
C 944 13 31 4 US-09-022-765-28 Sequence 28, Appl  
C 945 13 31 4 US-09-510-925A-1 Sequence 1, Appli  
C 946 13 33 4 US-09-238-356-33 Sequence 33, Appl  
C 947 13 34 2 US-08-846-762-40 Sequence 40, Appl  
C 948 13 34 3 US-09-133-914-1 Sequence 1, Appli  
C 949 13 34 4 US-09-469-197-1 Sequence 1, Appli  
C 950 13 36 2 US-08-816-155B-30 Sequence 30, Appl  
C 951 13 36 3 US-08-815-809-15 Sequence 15, Appl  
C 952 13 36 3 US-09-079-587-30 Sequence 30, Appl  
C 953 13 36 3 US-08-961-083-419 Sequence 419, App  
C 954 13 36 4 US-08-482-918-17 Sequence 17, Appl  
C 955 13 36 4 US-09-224-681-17 Sequence 17, Appl  
C 956 13 36 4 US-08-336-728A-17 Sequence 4, Appli  
C 957 13 37 2 US-08-898-560-4 Sequence 5, Appli  
C 958 13 37 2 US-08-898-560-5 Sequence 34, Appl  
959 13 37 4 US-09-518-914-34 Sequence 609, App  
960 13 38 1 US-08-373-124A-609 Sequence 1650, Ap  
961 13 38 1 US-08-373-124A-1650 Sequence 609, App  
962 13 38 1 US-08-435-628-609 Sequence 1650, Ap  
963 13 38 1 US-08-435-628-1650 Sequence 15, Appl  
C 964 13 38 4 US-09-110-517-15 Sequence 15, Appl  
965 13 39 2 US-08-951-871-25 Sequence 25, Appl  
966 13 42 4 US-09-498-959-5 Sequence 5, Appli  
967 13 44 2 US-08-875-154-28 Sequence 28, Appl  
C 968 13 49 4 US-09-538-709-1147 Sequence 1147, Ap  
969 13 55 4 US-09-345-882-6 Sequence 6, Appli  
970 13 56 1 US-07-910-760-3 Sequence 3, Appli  
971 13 56 1 US-08-444-818-241 Sequence 241, App  
972 13 56 3 US-08-440-519-3 Sequence 3, Appli  
973 13 56 4 US-08-440-549-3 Sequence 3, Appli  
974 13 56 5 PCT-US91-02225-3 Sequence 3, Appli  
C 975 13 60 4 US-08-654-737B-8 Sequence 8, Appli  
976 13 60 4 US-08-899-279-38 Sequence 38, Appl

977 13 1.3 60 4 US-08-899-279-38 Sequence 38, Appl  
C 978 13 1.3 61 1 US-08-451-472-28 Sequence 28, Appl  
979 13 1.3 61 1 US-08-451-472-29 Sequence 29, Appl  
C 980 13 1.3 73 4 US-08-952-793-168 Sequence 168, App  
C 981 13 1.3 73 5 PCT-US96-09455A-168 Sequence 168, App  
C 982 13 1.3 78 1 US-08-471-985A-106 Sequence 106, App  
C 983 13 1.3 78 5 PCT-US95-12401A-106 Sequence 106, App  
C 984 13 1.3 79 1 US-08-472-255A-136 Sequence 136, App  
C 985 13 1.3 79 1 US-08-479-724A-136 Sequence 136, App  
C 986 13 1.3 79 3 US-08-472-256B-136 Sequence 136, App  
C 987 13 1.3 79 4 US-08-687-421-399 Sequence 399, App  
C 988 13 1.3 79 4 US-08-952-793-136 Sequence 136, App  
C 989 13 1.3 79 5 PCT-US96-09455A-136 Sequence 136, App  
C 990 13 1.3 80 1 US-08-471-985A-51 Sequence 51, Appl  
C 991 13 1.3 80 1 US-08-471-985A-52 Sequence 52, Appl  
C 992 13 1.3 80 1 US-08-471-985A-83 Sequence 83, Appl  
C 993 13 1.3 80 1 US-08-472-255A-138 Sequence 138, App  
C 994 13 1.3 80 1 US-08-472-255A-149 Sequence 149, App  
C 995 13 1.3 80 1 US-08-472-255A-156 Sequence 156, App  
C 996 13 1.3 80 1 US-08-479-724A-138 Sequence 138, App  
C 997 13 1.3 80 1 US-08-479-724A-149 Sequence 149, App  
C 998 13 1.3 80 1 US-08-479-724A-156 Sequence 156, App  
C 999 13 1.3 80 3 US-08-472-256B-138 Sequence 138, App  
C1000 13 1.3 80 3 US-08-472-256B-149 Sequence 149, App

ALIGNMENTS

RESULT 1  
US-09-277-565-25  
; Sequence 25, Application US/09277565  
; Patent No. 6207384  
; GENERAL INFORMATION:  
; APPLICANT: Mekalanos, John J.  
; APPLICANT: Akerley, Brian J.  
; APPLICANT: Rubin, Eric J.  
; APPLICANT: Camilli, Andrew  
; TITLE OF INVENTION: SYSTEMATIC IDENTIFICATION OF ESSENTIAL  
; GENES BY IN VITRO TRANSPOSON MUTAGENESIS  
; FILE REFERENCE: 00742/052002  
; CURRENT APPLICATION NUMBER: US/09/277,565  
; CURRENT FILING DATE: 1999-03-26  
; EARLIER APPLICATION NUMBER: 60/079,770  
; EARLIER FILING DATE: 1998-03-27  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 25  
; LENGTH: 1728  
; TYPE: DNA  
; ORGANISM: Haemophilus influenzae  
US-09-277-565-25

Query Match 1.9%; Score 18; DB 4; Length 1728;  
Best Local Similarity 100.0%; Pred. No. 6.2;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 306 TGCTGCAAAAATGCCAA 323  
|  
Db 324 TGCTGCAAAAATGCCAA 341  
|

RESULT 2  
US-09-221-017B-356/c  
; Sequence 356, Application US/09221017B  
; Patent No. 6444799  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Bruce C.  
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
; NUMBER OF SEQUENCES: 1120  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD

CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/221,017B  
FILING DATE: 23-DEC-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1182  
FILING DATE: 31-DEC-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1546  
FILING DATE: 30-JAN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP2911  
FILING DATE: 09-APR-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU98/01023  
FILING DATE: 10-DEC-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Monroy, Gladys H  
REGISTRATION NUMBER: 32,430  
REFERENCE/DOCKET NUMBER: 27340-20021.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-813-5600  
TELEFAX: 650-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 356:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7930 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: UNKNOWN  
ORIGINAL SOURCE:  
ORGANISM: PORYPHYROMONAS GINGIVALIS  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1...7930  
US-09-221-017B-356

Query Match 1.9%; Score 18; DB 4; Length 7930;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 700 ATCATCAAGCGCAAAGGC 717  
|  
Db 5464 ATCATCAAGCGCAAAGGC 5447  
|

RESULT 3  
US-08-781-891-207/c  
; Sequence 207, Application US/08781891  
; Patent No. 6090620  
; GENERAL INFORMATION:  
; APPLICANT: Fu, Ying-Hui  
; APPLICANT: Yu, Chang-En  
; APPLICANT: Oshima, Junko  
; APPLICANT: Mulligan, John T.  
; APPLICANT: Schellenberg, Gerald D.  
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO  
; NUMBER OF SEQUENCES: 209  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781,891  
FILING DATE: 27-DEC-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 609062tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 240052.419  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 207:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29604 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-781-891-207

Query Match 1.9%; Score 18; DB 3; Length 29604;  
Best Local Similarity 100.0%; Pred.No. 7;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 311 CAAAAAATGCCAACATTA 328  
|||||  
Db 9553 CAAAAAATGCCAACATTA 9536

RESULT 4  
US-08-248-474-44  
Sequence 44, Application US/08248474  
Patent No. 5612471  
GENERAL INFORMATION:  
APPLICANT: MCK. BIRD, David  
APPLICANT: WILSON, Mark A.  
TITLE OF INVENTION: NEMATODE-INDUCE GENES IN TOMATO  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Khourie and Crew  
STREET: Stuart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/248,474  
FILING DATE: 25-MAY-1994  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 2307E-535  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 193 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Lycopersicon esculentum cv 'Rutgers Large Red'  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..193  
OTHER INFORMATION: /standard\_name= "DB# 168"  
US-08-248-474-44  
Query Match 1.7%; Score 17; DB 1; Length 193;  
Best Local Similarity 100.0%; Pred.No. 19;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 562 GATTCGGGTGTGCCTGT 578  
|||||  
Db 85 GATTCGGGTGTGCCTGT 101  
RESULT 5  
US-08-756-849-44  
Sequence 44, Application US/08756849  
Patent No. 6093810  
GENERAL INFORMATION:  
APPLICANT: Bird, David MCK.  
APPLICANT: Wilson, Mark A.  
TITLE OF INVENTION: Nematode-Induced Genes in Tomato  
NUMBER OF SEQUENCES: 129  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/756,849  
FILING DATE: 26-NOV-1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,474  
FILING DATE: 25-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 02307O-053510US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 193 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Lycopersicon esculentum cv 'Rutgers Large Red'  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..193  
OTHER INFORMATION: /standard\_name= "DB# 168"  
US-08-756-849-44  
Query Match 1.7%; Score 17; DB 3; Length 193;



Best Local Similarity 100.0%; Pred. No. 19;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 562 GATTCGGGTGTGCCTGT 578  
Db 85 GATTCGGGTGTGCCTGT 101  
  
RESULT 6  
US-09-315-794-61  
; Sequence 61, Application US/09315794  
; Patent No. 6197517  
; GENERAL INFORMATION:  
; APPLICANT: Roberts, Christopher J.  
; TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL  
; TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATION  
; TITLE OF INVENTION: DRUGS  
; FILE REFERENCE: 9301-053  
; CURRENT APPLICATION NUMBER: US/09/315,794  
; CURRENT FILING DATE: 1999-05-21  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 61  
; LENGTH: 402  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-09-315-794-61

Query Match 1.7%; Score 17; DB 4; Length 402;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 825 TGAACAAATTGCTATAA 841  
Db 15 TGAACAAATTGCTATAA 31

RESULT 7  
US-09-389-341-61  
; Sequence 61, Application US/09389341  
; Patent No. 6200803  
; GENERAL INFORMATION:  
; APPLICANT: Roberts, Christopher J.  
; TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL  
; TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATIVE  
; TITLE OF INVENTION: DRUGS  
; FILE REFERENCE: 9301-057  
; CURRENT APPLICATION NUMBER: US/09/389,341  
; CURRENT FILING DATE: 1999-09-02  
; EARLIER APPLICATION NUMBER: 09/315,794  
; EARLIER FILING DATE: 1999-05-21  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 61  
; LENGTH: 402  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-09-389-341-61

Query Match 1.7%; Score 17; DB 4; Length 402;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 825 TGAACAAATTGCTATAA 841  
Db 15 TGAACAAATTGCTATAA 31

RESULT 8  
US-09-173-300-16/c  
; Sequence 16, Application US/09173300  
; Patent No. 6451581  
; GENERAL INFORMATION:

; APPLICANT: Falco, Saverio Carl  
; APPLICANT: Hitz, William D.  
; APPLICANT: Kinney, Anthony J.  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Rafalski, J. Antoni  
; TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB-1126  
; CURRENT APPLICATION NUMBER: US/09/173,300  
; CURRENT FILING DATE: 1998-10-15  
; EARLIER APPLICATION NUMBER: 60/063,423  
; EARLIER FILING DATE: 1997 October 28  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: Microsoft Word Version 7.0A  
; SEQ ID NO 16  
; LENGTH: 965  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-173-300-16

Query Match 1.7%; Score 17; DB 4; Length 965;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 284 ATGGACAAACAAGGATG 300  
Db 17 ATGGACAAACAAGGATG 1

RESULT 9  
US-09-134-001C-1491  
; Sequence 1491, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 1491  
; LENGTH: 993  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-1491

Query Match 1.7%; Score 17; DB 4; Length 993;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 688 GCTGCTTACGATATCAT 704  
Db 700 GCTGCTTACGATATCAT 716

RESULT 10  
US-08-136-743B-1/c  
; Sequence 1, Application US/08136743B  
; Patent No. 5459063  
; GENERAL INFORMATION:  
; APPLICANT: Barry S. Cooperman, Harvey Rubin,  
; APPLICANT: Jerome Salem, and Alison L. Fisher  
; TITLE OF INVENTION: "Plasmodium falciparum Ribonu-  
; TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide Inhibit  
; TITLE OF INVENTION: Thereof"  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The University of Pennsylvania  
; STREET: Suite 330

STREET: 3700 Market Street  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19104-3246  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/136,743B  
FILING DATE: 10/14/93  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 3957-10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
TELEX: NO. 5459063e  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1112 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-136-743B-1

Query Match 1.7%; Score 17; DB 1; Length 1112;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 CTTGCTGCAAAAATGC 320  
|||||  
Db 272 CTTGCTGCAAAAATGC 256

RESULT 11  
US-08-869-506-1  
; Sequence 1, Application US/08869506  
; Patent No. 5827710  
; GENERAL INFORMATION:  
; APPLICANT: Uchida, Kohji  
; APPLICANT: Matsukawa, Hirokazu  
; APPLICANT: Matuo, Yushi  
; APPLICANT: Fujita, Tutosi  
; TITLE OF INVENTION: DNA ENCODING THE SUBUNIT OF AVIAN  
; TITLE OF INVENTION: LACTATE DEHYDROGENASE  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 No. 5827710th Glebe Rd. 8th floor  
; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201-4741  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/869,506  
; FILING DATE: 05-JUN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 73797/1996  
; FILING DATE: 29-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crawford, Arthur R.  
; REGISTRATION NUMBER: 25,327

REFERENCE/DOCKET NUMBER: 159-43  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1240 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 55..1053  
US-08-869-506-1

Query Match 1.7%; Score 17; DB 1; Length 1240;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 550 GGAGAACATGGTGATTC 566  
|||||  
Db 625 GGAGAACATGGTGATTC 641

RESULT 12  
US-09-128-967-1  
; Sequence 1, Application US/09128967  
; Patent No. 6057141  
; GENERAL INFORMATION:  
; APPLICANT: Uchida, Kohji  
; APPLICANT: Matsukawa, Hirokazu  
; APPLICANT: Matuo, Yushi  
; APPLICANT: Fujita, Tutosi  
; TITLE OF INVENTION: DNA ENCODING THE SUBUNIT OF AVIAN  
; TITLE OF INVENTION: LACTATE DEHYDROGENASE  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 No. 6057141th Glebe Rd. 8th floor  
; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201-4741  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/128,967  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/869,506  
; FILING DATE: 05-JUN-1997  
; APPLICATION NUMBER: JP 73797/1996  
; FILING DATE: 29-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crawford, Arthur R.  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 159-43  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1240 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:

```

; NAME/KEY: CDS
; LOCATION: 55..1053
US-09-128-967-1

Query Match 1.7%; Score 17; DB 3; Length 1240;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 550 GGAGAACATGGTGATTC 566
Db 625 GGAGAACATGGTGATTC 641

RESULT 13
US-09-234-186-2
; Sequence 2, Application US/09234186
; Patent No. 6312947
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Hengartner, Michael
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF A
; TITLE OF INVENTION: GENE WHICH PROTECTS CELLS FROM PROGRAMMED CELL DEATH AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 01997/201005
; CURRENT APPLICATION NUMBER: US/09/234,186
; CURRENT FILING DATE: 1999-01-20
; EARLIER APPLICATION NUMBER: 07/898,933
; EARLIER FILING DATE: 1992-06-12
; EARLIER APPLICATION NUMBER: 07/927,681
; EARLIER FILING DATE: 1992-08-10
; EARLIER APPLICATION NUMBER: 08/288,295
; EARLIER FILING DATE: 1994-08-10
; EARLIER APPLICATION NUMBER: 08/801,248
; EARLIER FILING DATE: 1997-02-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1315
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)...(846)
US-09-234-186-2

Query Match 1.7%; Score 17; DB 4; Length 1315;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 914 AATTGATGGAAAAATCA 930
Db 201 AATTGATGGAAAAATCA 217

RESULT 14
US-09-234-186-4
; Sequence 4, Application US/09234186
; Patent No. 6312947
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Hengartner, Michael
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF A
; TITLE OF INVENTION: GENE WHICH PROTECTS CELLS FROM PROGRAMMED CELL DEATH AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 01997/201005
; CURRENT APPLICATION NUMBER: US/09/234,186
; CURRENT FILING DATE: 1999-01-20
; EARLIER APPLICATION NUMBER: 07/898,933
; EARLIER FILING DATE: 1992-06-12
; EARLIER APPLICATION NUMBER: 07/927,681
; EARLIER FILING DATE: 1992-08-10
; EARLIER APPLICATION NUMBER: 08/288,295
; EARLIER FILING DATE: 1994-08-10

```

```

; EARLIER APPLICATION NUMBER: 08/801,248
; EARLIER FILING DATE: 1997-02-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1315
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)...(846)
US-09-234-186-4

Query Match 1.7%; Score 17; DB 4; Length 1315;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 914 AATTGATGGAAAAATCA 930
Db 201 AATTGATGGAAAAATCA 217

RESULT 15
US-09-234-186-5
; Sequence 5, Application US/09234186
; Patent No. 6312947
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Hengartner, Michael
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF A
; TITLE OF INVENTION: GENE WHICH PROTECTS CELLS FROM PROGRAMMED CELL DEATH AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 01997/201005
; CURRENT APPLICATION NUMBER: US/09/234,186
; CURRENT FILING DATE: 1999-01-20
; EARLIER APPLICATION NUMBER: 07/898,933
; EARLIER FILING DATE: 1992-06-12
; EARLIER APPLICATION NUMBER: 07/927,681
; EARLIER FILING DATE: 1992-08-10
; EARLIER APPLICATION NUMBER: 08/288,295
; EARLIER FILING DATE: 1994-08-10
; EARLIER APPLICATION NUMBER: 08/801,248
; EARLIER FILING DATE: 1997-02-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1315
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)...(846)
US-09-234-186-5

Query Match 1.7%; Score 17; DB 4; Length 1315;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 914 AATTGATGGAAAAATCA 930
Db 201 AATTGATGGAAAAATCA 217

Search completed: July 31, 2003, 12:44:40
Job time : 98 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 11:59:02 ; Search time 260 Seconds  
(without alignments)  
7712.449 Million cell updates/sec

Title: US-09-992-430B-21  
Perfect score: 972  
Sequence: 1 atgttccaagatacaaaagtc.....ttgagcatctggagatcaat 972

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0  
Searched: 1439767 seqs, 1031500376 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq1:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq3:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	2.0	502	12	US-09-918-995-24455
2	19	2.0	530	11	US-09-833-381-773
3	19	2.0	722	15	US-10-027-632-124906
4	19	2.0	1146	15	US-10-274-266-1
5	19	2.0	1755	11	US-09-731-872-15
6	19	2.0	1759	11	US-09-731-872-14
7	19	2.0	2352	15	US-10-128-714-87
8	19	2.0	2352	15	US-10-128-714-5087
9	19	2.0	3144	15	US-10-274-266-3
10	18	1.9	492	12	US-09-918-995-11651
11	18	1.9	559	15	US-10-027-632-233560
12	18	1.9	828	11	US-09-974-300-6602
13	18	1.9	996	11	US-09-938-842A-4346
14	18	1.9	996	11	US-09-938-842A-5127
15	18	1.9	2338	15	US-10-027-632-259728
16	18	1.9	2338	15	US-10-027-632-259729

17	18	1.9	2338	15	US-10-027-632-259730	Sequence 259730,
18	18	1.9	1830121	15	US-10-329-960-1	Sequence 1, Appli
19	17	1.7	98	10	US-09-864-761-26075	Sequence 26075, A
20	17	1.7	329	11	US-09-783-590-1290	Sequence 1290, Ap
21	17	1.7	344	10	US-09-777-564-756	Sequence 756, App
22	17	1.7	344	10	US-09-777-564-1494	Sequence 1494, Ap
23	17	1.7	344	15	US-10-015-219-756	Sequence 756, App
24	17	1.7	344	15	US-10-015-219-1494	Sequence 1494, Ap
25	17	1.7	377	11	US-09-983-965-5432	Sequence 5432, Ap
26	17	1.7	422	12	US-09-918-995-5225	Sequence 5225, Ap
27	17	1.7	432	11	US-09-783-590-6454	Sequence 6454, Ap
28	17	1.7	487	12	US-09-918-995-12259	Sequence 12259, A
29	17	1.7	542	15	US-10-027-632-91808	Sequence 91808, A
30	17	1.7	542	15	US-10-027-632-304597	Sequence 304597,
31	17	1.7	559	15	US-10-027-632-202059	Sequence 202059,
32	17	1.7	569	10	US-09-864-761-9660	Sequence 9660, Ap
33	17	1.7	572	15	US-10-027-632-217575	Sequence 217575,
34	17	1.7	598	15	US-10-027-632-300433	Sequence 300433,
35	17	1.7	598	15	US-10-027-632-300434	Sequence 300434,
36	17	1.7	609	15	US-10-027-632-114606	Sequence 114606,
37	17	1.7	612	15	US-10-027-632-77370	Sequence 77370, A
38	17	1.7	640	15	US-10-027-632-113106	Sequence 113106,
39	17	1.7	647	15	US-10-027-632-115362	Sequence 115362,
40	17	1.7	965	15	US-10-027-450-16	Sequence 16, Appl
41	17	1.7	971	15	US-10-027-632-121628	Sequence 121628,
42	17	1.7	971	15	US-10-027-632-121629	Sequence 121629,
43	17	1.7	971	15	US-10-027-632-121630	Sequence 121630,
44	17	1.7	1009	12	US-09-934-455-507	Sequence 507, App
45	17	1.7	1344	11	US-09-738-626-3332	Sequence 3332, Ap
46	17	1.7	1404	10	US-09-765-272-9	Sequence 9, Appli
47	17	1.7	1657	15	US-10-027-632-255088	Sequence 255088,
48	17	1.7	1657	15	US-10-027-632-255089	Sequence 255089,
49	17	1.7	1884	11	US-09-731-872-18	Sequence 18, Appl
50	17	1.7	1992	11	US-09-917-800A-1716	Sequence 1716, Ap
51	17	1.7	3129	15	US-10-161-803-24	Sequence 24, Appl
52	17	1.7	3172	15	US-10-013-477-6	Sequence 6, Appli
53	17	1.7	3799	10	US-09-866-562-24	Sequence 24, Appl
54	17	1.7	3799	11	US-09-960-253-161	Sequence 161, App
55	17	1.7	4082	11	US-09-864-864-325	Sequence 325, App
56	17	1.7	4240	12	US-09-919-039-208	Sequence 208, App
57	17	1.7	6765	15	US-10-037-270-481	Sequence 481, App
58	17	1.7	14962	10	US-09-764-878-244	Sequence 244, App
59	17	1.7	14962	15	US-10-079-854-244	Sequence 244, App
60	17	1.7	20099	15	US-10-177-744A-12	Sequence 12, Appl
61	17	1.7	21724	11	US-09-764-864-1603	Sequence 1603, Ap
62	17	1.7	21727	11	US-09-764-864-1604	Sequence 1604, Ap
63	17	1.7	536165	12	US-09-939-964-1	Sequence 1, Appli
64	17	1.7	3309400	11	US-09-738-626-1	Sequence 1, Appli
65	16	1.6	25	15	US-10-098-263B-86358	Sequence 86358, A
66	16	1.6	81	12	US-09-849-928-166	Sequence 166, App
67	16	1.6	81	15	US-10-066-960-166	Sequence 166, App
68	16	1.6	129	15	US-10-300-285-7	Sequence 7, Appli
69	16	1.6	129	15	US-10-300-285-8	Sequence 8, Appli
70	16	1.6	189	12	US-09-935-916A-27	Sequence 27, Appl
71	16	1.6	189	12	US-09-935-916A-29	Sequence 29, Appl
72	16	1.6	189	12	US-09-935-916A-30	Sequence 30, Appl
73	16	1.6	189	12	US-09-935-916A-31	Sequence 31, Appl
74	16	1.6	189	12	US-09-935-916A-33	Sequence 33, Appl
75	16	1.6	204	11	US-09-878-574-7747	Sequence 7747, Ap
76	16	1.6	240	10	US-09-864-761-29924	Sequence 29924, A
77	16	1.6	240	11	US-09-783-590-7301	Sequence 7301, Ap
78	16	1.6	262	11	US-09-878-574-10105	Sequence 10105, A
79	16	1.6	265	11	US-09-878-574-453	Sequence 453, App
80	16	1.6	295	11	US-09-796-692-4124	Sequence 4124, Ap
81	16	1.6	295	15	US-10-040-862-4124	Sequence 4124, Ap
82	16	1.6	359	11	US-09-878-574-3321	Sequence 3321, Ap
83	16	1.6	403	12	US-09-918-995-18597	Sequence 18597, A
84	16	1.6	406	12	US-09-918-995-4285	Sequence 4285, Ap
85	16	1.6	409	11	US-09-796-692-9430	Sequence 9430, Ap
86	16	1.6	409	15	US-10-040-862-9430	Sequence 9430, Ap
87	16	1.6	418	11	US-09-983-965-5780	Sequence 5780, Ap
88	16	1.6	426	15	US-10-106-698-1747	Sequence 1747, Ap
89	16	1.6	427	12	US-09-918-995-1646	Sequence 1646, Ap



c 90	16	1.6	438	12	US-09-991-936-258	Sequence 258, App	163	16	1.6	829	11	US-09-070-927A-686	Sequence 686, App
c 91	16	1.6	446	15	US-10-171-581-315	Sequence 315, App	164	16	1.6	878	15	US-10-027-632-122083	Sequence 122083,
c 92	16	1.6	450	11	US-09-925-300-939	Sequence 939, App	165	16	1.6	960	10	US-09-939-980-46	Sequence 46, Appl
c 93	16	1.6	452	15	US-10-040-739-1119	Sequence 1119, Ap	166	16	1.6	978	12	US-09-991-936-16	Sequence 16, Appl
c 94	16	1.6	462	10	US-09-864-761-14088	Sequence 14088, A	c 167	16	1.6	978	12	US-09-991-936-18	Sequence 18, Appl
c 95	16	1.6	464	12	US-09-918-995-15685	Sequence 15685, A	c 168	16	1.6	987	10	US-09-815-242-9774	Sequence 9774, Ap
c 96	16	1.6	474	15	US-10-027-632-192630	Sequence 192630,	169	16	1.6	990	15	US-10-208-018-103	Sequence 103, App
c 97	16	1.6	482	12	US-09-918-995-12096	Sequence 12096, A	170	16	1.6	1036	11	US-09-925-300-713	Sequence 713, App
c 98	16	1.6	485	15	US-10-027-632-195557	Sequence 195557,	c 171	16	1.6	1039	15	US-10-071-766-41	Sequence 41, Appl
c 99	16	1.6	485	15	US-10-027-632-195558	Sequence 195558,	172	16	1.6	1066	10	US-09-768-826-20	Sequence 20, Appl
c 100	16	1.6	485	15	US-10-027-632-195559	Sequence 195559,	173	16	1.6	1068	11	US-09-974-300-739	Sequence 739, App
c 101	16	1.6	487	15	US-10-027-632-290017	Sequence 290017,	174	16	1.6	1069	15	US-10-027-632-252428	Sequence 252428,
c 102	16	1.6	488	15	US-10-027-632-38029	Sequence 38029, A	175	16	1.6	1219	15	US-10-027-632-241194	Sequence 241194,
c 103	16	1.6	492	12	US-09-918-995-30493	Sequence 30493, A	176	16	1.6	1239	10	US-09-815-242-7880	Sequence 7880, Ap
c 104	16	1.6	495	12	US-09-764-891-6606	Sequence 6606, Ap	c 177	16	1.6	1299	11	US-09-974-300-674	Sequence 674, App
c 105	16	1.6	495	15	US-10-027-632-215369	Sequence 215369,	178	16	1.6	1345	10	US-09-864-761-19745	Sequence 19745, A
c 106	16	1.6	495	15	US-10-091-572-470	Sequence 470, App	c 179	16	1.6	1371	11	US-09-801-368-203	Sequence 203, App
c 107	16	1.6	525	15	US-10-027-632-84506	Sequence 84506, A	c 180	16	1.6	1476	11	US-09-801-368-77	Sequence 77, Appl
c 108	16	1.6	527	11	US-09-833-381-1083	Sequence 1083, Ap	c 181	16	1.6	1484	11	US-09-764-868-429	Sequence 429, App
c 109	16	1.6	528	15	US-10-060-036-3382	Sequence 3382, Ap	c 182	16	1.6	1521	12	US-09-984-827-12	Sequence 12, Appl
c 110	16	1.6	546	15	US-10-027-632-180787	Sequence 180787,	183	16	1.6	1536	12	US-09-918-624B-37	Sequence 37, Appl
c 111	16	1.6	554	10	US-09-864-761-13363	Sequence 13363, A	c 184	16	1.6	1683	10	US-09-815-242-7649	Sequence 7649, Ap
c 112	16	1.6	559	15	US-10-027-632-293804	Sequence 293804,	185	16	1.6	1714	12	US-09-991-936-13	Sequence 13, Appl
c 113	16	1.6	559	15	US-10-027-632-321524	Sequence 321524,	c 186	16	1.6	1714	12	US-09-991-936-15	Sequence 15, Appl
c 114	16	1.6	560	15	US-10-027-632-48209	Sequence 48209, A	187	16	1.6	1788	10	US-09-815-242-4189	Sequence 4189, Ap
c 115	16	1.6	560	15	US-10-027-632-48210	Sequence 48210, A	188	16	1.6	1815	10	US-09-815-242-8427	Sequence 8427, Ap
c 116	16	1.6	560	15	US-10-027-632-48211	Sequence 48211, A	189	16	1.6	1833	11	US-09-938-842A-2397	Sequence 2397, Ap
c 117	16	1.6	575	15	US-10-027-632-292018	Sequence 290018,	190	16	1.6	1854	15	US-10-037-270-1060	Sequence 1060, Ap
c 118	16	1.6	577	15	US-10-027-632-73404	Sequence 73404, A	191	16	1.6	1967	10	US-09-864-761-2964	Sequence 2964, Ap
c 119	16	1.6	577	15	US-10-027-632-73405	Sequence 73405, A	192	16	1.6	1967	15	US-10-023-896-20	Sequence 20, Appl
c 120	16	1.6	577	15	US-10-027-632-73406	Sequence 73406, A	c 193	16	1.6	2000	11	US-09-938-842A-3721	Sequence 3721, Ap
c 121	16	1.6	577	15	US-10-027-632-292810	Sequence 292810,	194	16	1.6	2000	11	US-09-938-842A-4851	Sequence 4851, Ap
c 122	16	1.6	577	15	US-10-027-632-292811	Sequence 292811,	c 195	16	1.6	2000	11	US-09-938-842A-5195	Sequence 5195, Ap
c 123	16	1.6	577	15	US-10-027-632-292812	Sequence 292812,	196	16	1.6	2063	15	US-10-051-909-33	Sequence 33, Appl
c 124	16	1.6	580	12	US-09-764-891-2558	Sequence 2558, Ap	197	16	1.6	2125	10	US-09-880-192-34	Sequence 34, Appl
c 125	16	1.6	595	15	US-10-103-313-20	Sequence 20, Appl	198	16	1.6	2147	11	US-09-764-855-326	Sequence 326, App
c 126	16	1.6	603	15	US-10-027-632-227134	Sequence 227134,	199	16	1.6	2147	15	US-10-072-349-326	Sequence 326, App
c 127	16	1.6	603	15	US-10-027-632-227135	Sequence 227135,	c 200	16	1.6	2190	15	US-10-027-632-261915	Sequence 261915,
c 128	16	1.6	611	15	US-10-027-632-203536	Sequence 203536,	201	16	1.6	2236	12	US-09-764-891-6383	Sequence 6383, Ap
c 129	16	1.6	615	15	US-10-027-632-238921	Sequence 238921,	202	16	1.6	2236	15	US-10-205-428-673	Sequence 673, Ap
c 130	16	1.6	618	15	US-10-027-632-198651	Sequence 198651,	c 203	16	1.6	2266	11	US-09-764-864-17	Sequence 17, Appl
c 131	16	1.6	623	15	US-10-119-926-88	Sequence 88, Appl	204	16	1.6	2268	10	US-09-815-242-6058	Sequence 6058, Ap
c 132	16	1.6	624	15	US-10-027-632-228376	Sequence 228376,	c 205	16	1.6	2344	15	US-10-024-632-1	Sequence 1, Appli
c 133	16	1.6	624	15	US-10-027-632-228377	Sequence 228377,	206	16	1.6	2394	10	US-09-349-015-31	Sequence 31, Appl
c 134	16	1.6	626	15	US-10-027-632-114107	Sequence 114107,	c 207	16	1.6	2619	10	US-09-925-301-162	Sequence 162, App
c 135	16	1.6	632	15	US-10-027-632-221715	Sequence 221715,	c 208	16	1.6	2702	11	US-09-834-975-837	Sequence 837, App
c 136	16	1.6	634	15	US-10-027-632-252572	Sequence 252572,	c 209	16	1.6	2702	11	US-09-834-975-841	Sequence 841, App
c 137	16	1.6	637	15	US-10-027-632-195859	Sequence 195859,	c 210	16	1.6	2790	15	US-10-072-094-98	Sequence 98, Appl
c 138	16	1.6	637	15	US-10-027-632-220926	Sequence 220926,	c 211	16	1.6	2844	15	US-10-044-090-813	Sequence 813, App
c 139	16	1.6	645	12	US-09-764-891-6671	Sequence 6671, Ap	c 212	16	1.6	2890	15	US-10-027-632-256651	Sequence 256651,
c 140	16	1.6	645	12	US-09-764-891-6674	Sequence 6674, Ap	c 213	16	1.6	2991	15	US-10-201-386-48	Sequence 48, Appl
c 141	16	1.6	645	15	US-10-091-572-532	Sequence 532, App	c 214	16	1.6	3168	15	US-10-027-632-114963	Sequence 114963,
c 142	16	1.6	645	15	US-10-091-572-535	Sequence 535, App	c 215	16	1.6	3367	15	US-10-173-539-7	Sequence 7, Appli
c 143	16	1.6	647	15	US-10-027-632-44349	Sequence 44349, A	c 216	16	1.6	3368	15	US-10-198-846-9722	Sequence 9722, Ap
c 144	16	1.6	647	15	US-10-027-632-244428	Sequence 244428,	c 217	16	1.6	3499	15	US-10-173-539-3	Sequence 3, Appli
c 145	16	1.6	647	15	US-10-027-632-244429	Sequence 244429,	c 218	16	1.6	3550	15	US-10-173-539-13	Sequence 13, Appl
c 146	16	1.6	661	15	US-10-027-632-287226	Sequence 287226,	c 219	16	1.6	3824	15	US-10-125-792-11	Sequence 11, Appl
c 147	16	1.6	661	15	US-10-027-632-287227	Sequence 287227,	c 220	16	1.6	3824	15	US-10-125-778-11	Sequence 11, Appl
c 148	16	1.6	666	15	US-10-027-632-270003	Sequence 270003,	c 221	16	1.6	3824	15	US-10-125-772-11	Sequence 11, Appl
c 149	16	1.6	676	15	US-10-027-632-205278	Sequence 205278,	c 222	16	1.6	3941	15	US-10-125-792-7	Sequence 7, Appli
c 150	16	1.6	697	15	US-10-027-632-286967	Sequence 286967,	c 223	16	1.6	3941	15	US-10-125-778-7	Sequence 7, Appli
c 151	16	1.6	702	15	US-10-071-766-40	Sequence 40, Appl	c 224	16	1.6	3941	15	US-10-125-772-7	Sequence 7, Appli
c 152	16	1.6	715	15	US-10-027-632-28353	Sequence 28353, A	c 225	16	1.6	4031	15	US-10-125-792-9	Sequence 9, Appli
c 153	16	1.6	715	15	US-10-027-632-28354	Sequence 28354, A	c 226	16	1.6	4031	15	US-10-125-778-9	Sequence 9, Appli
c 154	16	1.6	715	15	US-10-027-632-28355	Sequence 28355, A	c 227	16	1.6	4031	15	US-10-125-772-9	Sequence 9, Appli
c 155	16	1.6	720	15	US-10-027-632-27106	Sequence 27106, A	c 228	16	1.6	4039	15	US-10-097-340-84	Sequence 84, Appl
c 156	16	1.6	720	15	US-10-027-632-27107	Sequence 27107, A	c 229	16	1.6	4154	15	US-10-182-232-5	Sequence 5, Appli
c 157	16	1.6	755	15	US-10-027-632-148827	Sequence 148827,	c 230	16	1.6	4302	11	US-09-801-368-59	Sequence 59, Appl
c 158	16	1.6	791	11	US-09-989-919-48	Sequence 48, Appl	c 231	16	1.6	5311	12	US-09-764-891-10174	Sequence 10174, A
c 159	16	1.6	798	15	US-10-027-632-28904	Sequence 28904, A	c 232	16	1.6	6651	11	US-09-954-456-178	Sequence 178, App
c 160	16	1.6	818	15	US-10-027-632-167384	Sequence 167384,	c 233	16	1.6	6651	11	US-09-954-456-1217	Sequence 1217, Ap
c 161	16	1.6	819	15	US-10-027-632-159384	Sequence 159384,	234	16	1.6	6651	15	US-10-205-823-92	Sequence 92, Appl
c 162	16	1.6	825	11	US-09-738-626-994	Sequence 994, App	235	16	1.6	6763	15	US-10-176-847-55	Sequence 55, Appl

C 236	16	1.6	7303	15	US-10-173-539-15	Sequence 15, Appl	Sequence 15, Appl	309	15	1.5	375	10	US-09-923-779-37	Sequence 37, Appl
C 237	16	1.6	8473	10	US-09-851-682A-2	Sequence 2, Appli	Sequence 2, Appli	310	15	1.5	375	10	US-09-923-779-116	Sequence 116, App
C 238	16	1.6	10236	15	US-10-114-170-240	Sequence 240, App	Sequence 240, App	C 311	15	1.5	375	10	US-09-923-779-117	Sequence 117, App
C 239	16	1.6	15231	11	US-09-917-800A-1505	Sequence 1505, Ap	Sequence 1505, Ap	312	15	1.5	375	10	US-09-923-779-128	Sequence 128, App
C 240	16	1.6	15366	10	US-09-764-860-1057	Sequence 1057, Ap	Sequence 1057, Ap	C 313	15	1.5	375	10	US-09-923-779-129	Sequence 129, App
C 241	16	1.6	15366	15	US-10-074-095-1057	Sequence 1057, Ap	Sequence 1057, Ap	C 314	15	1.5	375	10	US-09-923-779-141	Sequence 141, App
C 242	16	1.6	18449	11	US-09-764-868-1457	Sequence 1457, Ap	Sequence 1457, Ap	C 315	15	1.5	377	11	US-09-954-456-232	Sequence 232, App
C 243	16	1.6	21423	11	US-09-764-877-2835	Sequence 2835, Ap	Sequence 2835, Ap	316	15	1.5	377	15	US-10-027-632-253326	Sequence 253326,
C 244	16	1.6	31208	10	US-09-852-067-3	Sequence 3, Appli	Sequence 3, Appli	C 317	15	1.5	385	11	US-09-878-574-760	Sequence 760, App
C 245	16	1.6	31208	15	US-10-338-691-3	Sequence 3, Appli	Sequence 3, Appli	C 318	15	1.5	387	12	US-09-764-891-8049	Sequence 8049, Ap
C 246	16	1.6	32190	10	US-09-764-878-201	Sequence 201, App	Sequence 201, App	C 319	15	1.5	387	12	US-09-764-891-8050	Sequence 8050, Ap
C 247	16	1.6	32190	15	US-10-079-854-201	Sequence 201, App	Sequence 201, App	C 320	15	1.5	387	12	US-09-764-891-8051	Sequence 8051, Ap
C 248	16	1.6	32193	10	US-09-764-878-200	Sequence 200, App	Sequence 200, App	321	15	1.5	396	10	US-09-825-294-138	Sequence 138, App
C 249	16	1.6	32193	15	US-10-079-854-200	Sequence 200, App	Sequence 200, App	322	15	1.5	396	11	US-09-970-966-138	Sequence 138, App
C 250	16	1.6	32249	10	US-09-764-878-202	Sequence 202, App	Sequence 202, App	323	15	1.5	396	14	US-10-212-677-138	Sequence 138, App
C 251	16	1.6	32249	15	US-10-079-854-202	Sequence 202, App	Sequence 202, App	C 324	15	1.5	400	8	US-08-781-986A-1502	Sequence 1502, Ap
C 252	16	1.6	43058	11	US-09-954-456-292	Sequence 292, App	Sequence 292, App	C 325	15	1.5	400	8	US-08-781-986A-1502	Sequence 1502, Ap
C 253	16	1.6	43058	11	US-09-954-456-529	Sequence 529, App	Sequence 529, App	C 326	15	1.5	401	10	US-09-922-217-45	Sequence 4032, Ap
C 254	16	1.6	43058	11	US-09-880-107-3950	Sequence 3950, Ap	Sequence 3950, Ap	327	15	1.5	401	11	US-09-833-263-45	Sequence 45, Appl
C 255	16	1.6	55795	11	US-09-880-107-1543	Sequence 1543, Ap	Sequence 1543, Ap	328	15	1.5	401	15	US-10-025-380-45	Sequence 45, Appl
C 256	16	1.6	145831	11	US-09-969-708-79	Sequence 79, Appl	Sequence 79, Appl	C 329	15	1.5	405	11	US-10-025-380-45	Sequence 45, Appl
C 257	16	1.6	145831	11	US-09-954-456-2116	Sequence 2116, Ap	Sequence 2116, Ap	C 330	15	1.5	405	11	US-09-560-863-492	Sequence 492, App
C 258	16	1.6	402850	12	US-09-844-653-5	Sequence 5, Appli	Sequence 5, Appli	C 331	15	1.5	407	15	US-10-027-632-140856	Sequence 140856,
C 259	16	1.6	1503841	10	US-09-795-668-1	Sequence 1, Appli	Sequence 1, Appli	332	15	1.5	408	15	US-10-198-846-12667	Sequence 12667, A
C 260	16	1.6	1503841	10	US-09-795-686-1	Sequence 1, Appli	Sequence 1, Appli	333	15	1.5	410	11	US-09-960-352-4202	Sequence 4202, Ap
C 261	16	1.6	1503841	11	US-09-946-807-1	Sequence 1, Appli	Sequence 1, Appli	C 334	15	1.5	418	10	US-09-962-436-220	Sequence 220, App
C 262	16	1.6	1691139	15	US-10-067-514-1	Sequence 1, Appli	Sequence 1, Appli	C 335	15	1.5	418	12	US-09-918-995-1055	Sequence 1055, Ap
C 263	16	1.6	3309400	11	US-09-738-626-1	Sequence 1, Appli	Sequence 1, Appli	C 336	15	1.5	420	12	US-09-991-936-362	Sequence 362, App
C 264	15	1.5	24	11	US-09-798-584-10	Sequence 10, Appl	Sequence 10, Appl	337	15	1.5	420	15	US-10-027-632-183086	Sequence 183086,
C 265	15	1.5	25	15	US-10-098-263B-69065	Sequence 69065, A	Sequence 69065, A	338	15	1.5	422	12	US-09-918-995-8357	Sequence 8357, Ap
C 266	15	1.5	39	15	US-10-118-079-16	Sequence 16, Appl	Sequence 16, Appl	C 339	15	1.5	425	15	US-10-027-632-183752	Sequence 183752,
C 267	15	1.5	45	12	US-09-940-244-158	Sequence 158, App	Sequence 158, App	C 340	15	1.5	428	11	US-09-960-352-601	Sequence 601, App
C 268	15	1.5	45	12	US-09-864-636A-592	Sequence 592, App	Sequence 592, App	C 341	15	1.5	430	10	US-09-864-761-14096	Sequence 14096, A
C 269	15	1.5	45	15	US-10-033-297-158	Sequence 158, App	Sequence 158, App	C 342	15	1.5	430	12	US-09-918-995-32630	Sequence 32630, A
C 270	15	1.5	155	11	US-09-764-877-879	Sequence 879, App	Sequence 879, App	C 343	15	1.5	436	10	US-09-918-686-24	Sequence 24, Appl
C 271	15	1.5	173	11	US-09-736-457-503	Sequence 503, App	Sequence 503, App	C 344	15	1.5	439	11	US-09-960-352-3931	Sequence 3931, Ap
C 272	15	1.5	173	11	US-09-902-941-503	Sequence 503, App	Sequence 503, App	C 345	15	1.5	440	15	US-10-027-632-42837	Sequence 42837, A
C 273	15	1.5	173	11	US-09-849-626-503	Sequence 503, App	Sequence 503, App	C 346	15	1.5	446	15	US-10-027-632-62097	Sequence 62097, A
C 274	15	1.5	173	12	US-09-476-300-503	Sequence 503, App	Sequence 503, App	C 347	15	1.5	446	15	US-10-027-632-63418	Sequence 63418, A
C 275	15	1.5	173	15	US-10-017-754-503	Sequence 503, App	Sequence 503, App	C 348	15	1.5	447	15	US-10-027-632-297939	Sequence 297939,
C 276	15	1.5	188	10	US-09-912-020-77	Sequence 77, Appl	Sequence 77, Appl	349	15	1.5	450	12	US-09-738-626-2797	Sequence 2797, Ap
C 277	15	1.5	211	11	US-09-983-965-5235	Sequence 5235, Ap	Sequence 5235, Ap	C 350	15	1.5	450	12	US-09-918-995-25578	Sequence 25578, A
C 278	15	1.5	214	11	US-09-960-352-8594	Sequence 8594, Ap	Sequence 8594, Ap	C 351	15	1.5	454	15	US-10-027-632-87288	Sequence 87288, A
C 280	15	1.5	215	10	US-09-923-876-4635	Sequence 4635, Ap	Sequence 4635, Ap	C 352	15	1.5	454	15	US-10-027-632-182405	Sequence 182405,
C 281	15	1.5	232	10	US-09-923-876-3597	Sequence 3597, Ap	Sequence 3597, Ap	C 353	15	1.5	455	10	US-09-922-217-383	Sequence 383, App
C 282	15	1.5	233	11	US-09-878-574-7773	Sequence 7773, Ap	Sequence 7773, Ap	C 354	15	1.5	455	15	US-09-833-263-383	Sequence 383, App
C 283	15	1.5	245	15	US-10-027-632-275440	Sequence 275440,	Sequence 275440,	C 355	15	1.5	455	15	US-10-025-380-383	Sequence 181018,
C 284	15	1.5	247	11	US-09-878-574-10727	Sequence 10727, A	Sequence 10727, A	C 356	15	1.5	456	10	US-10-027-632-14332	Sequence 14332, A
C 285	15	1.5	258	11	US-09-878-574-12810	Sequence 12810, A	Sequence 12810, A	C 357	15	1.5	456	12	US-09-918-995-27800	Sequence 27800, A
C 286	15	1.5	267	12	US-09-803-719-1647	Sequence 1647, Ap	Sequence 1647, Ap	C 358	15	1.5	457	11	US-09-983-965-178	Sequence 178, App
C 287	15	1.5	269	10	US-09-294-093B-3909	Sequence 3909, Ap	Sequence 3909, Ap	359	15	1.5	461	12	US-09-918-995-2358	Sequence 2358, Ap
C 288	15	1.5	275	11	US-09-878-574-14072	Sequence 14072, A	Sequence 14072, A	C 360	15	1.5	461	15	US-10-027-632-274412	Sequence 274412,
C 289	15	1.5	276	11	US-09-738-626-2854	Sequence 2854, Ap	Sequence 2854, Ap	C 361	15	1.5	462	12	US-09-918-995-32802	Sequence 32802, A
C 290	15	1.5	284	12	US-09-232-785-244	Sequence 244, App	Sequence 244, App	C 362	15	1.5	462	15	US-09-918-995-32802	Sequence 32802, A
C 291	15	1.5	287	10	US-09-294-093B-3148	Sequence 3148, Ap	Sequence 3148, Ap	C 363	15	1.5	463	15	US-10-027-632-183930	Sequence 183930,
C 292	15	1.5	302	11	US-09-960-352-2799	Sequence 2799, Ap	Sequence 2799, Ap	364	15	1.5	464	15	US-10-027-632-183931	Sequence 183931,
C 293	15	1.5	303	10	US-09-741-669-130	Sequence 130, App	Sequence 130, App	365	15	1.5	464	15	US-10-027-632-269826	Sequence 269826,
C 294	15	1.5	303	10	US-09-912-020-230	Sequence 230, App	Sequence 230, App	C 366	15	1.5	467	11	US-10-027-632-269827	Sequence 269827,
C 295	15	1.5	303	10	US-09-815-242-6260	Sequence 6260, Ap	Sequence 6260, Ap	C 367	15	1.5	467	15	US-09-920-300A-911	Sequence 911, App
C 296	15	1.5	309	11	US-09-880-107-452	Sequence 452, App	Sequence 452, App	C 368	15	1.5	467	15	US-10-033-528-911	Sequence 911, App
C 297	15	1.5	321	12	US-09-764-891-1541	Sequence 1541, Ap	Sequence 1541, Ap	C 369	15	1.5	467	15	US-10-027-632-214490	Sequence 214490,
C 298	15	1.5	345	12	US-09-910-082A-28	Sequence 28, Appl	Sequence 28, Appl	C 370	15	1.5	467	15	US-10-027-632-274411	Sequence 274411,
C 299	15	1.5	351	10	US-09-745-288-77	Sequence 77, Appl	Sequence 77, Appl	C 371	15	1.5	467	15	US-10-027-632-274413	Sequence 274413,
C 300	15	1.5	351	11	US-09-974-300-3760	Sequence 3760, Ap	Sequence 3760, Ap	C 372	15	1.5	470	10	US-09-867-550-1295	Sequence 1295, Ap
C 301	15	1.5	354	11	US-09-738-626-2743	Sequence 2743, Ap	Sequence 2743, Ap	373	15	1.5	470	15	US-10-027-632-278381	Sequence 278381,
C 302	15	1.5	361	12	US-09-918-995-7584	Sequence 7584, Ap	Sequence 7584, Ap	374	15	1.5	470	15	US-10-027-632-278382	Sequence 278382,
C 303	15	1.5	362	15	US-10-198-846-12626	Sequence 12626, A	Sequence 12626, A	375	15	1.5	470	15	US-10-027-632-278383	Sequence 278383,
C 304	15	1.5	363	10	US-09-923-779-111	Sequence 111, App	Sequence 111, App	376	15	1.5	472	11	US-09-998-598-678	Sequence 678, App
C 305	15	1.5	363	10	US-09-923-779-112	Sequence 112, App	Sequence 112, App	377	15	1.5	472	12	US-09-918-995-22192	Sequence 22192, A
C 306	15	1.5	366	12	US-09-918-995-30171	Sequence 30171, A	Sequence 30171, A	378	15	1.5	475	11	US-09-764-847-413	Sequence 413, App
C 307	15	1.5	368	12	US-09-918-995-30102	Sequence 30102, A	Sequence 30102, A	379	15	1.5	475	15	US-10-092-154-413	Sequence 413, App
C 308	15	1.5	373	11	US-09-867-701-2195	Sequence 2195, Ap	Sequence 2195, Ap	C 380	15	1.5	476	12	US-09-918-995-14319	Sequence 14319, A
								C 381	15	1.5	480	11	US-09-924-035A-374	Sequence 374, App

C 382	15	1.5	481	10	US-09-864-761-14464	Sequence 14464, A	C 455	15	1.5	568	15	US-10-027-632-73890	Sequence 73890, A
C 383	15	1.5	481	15	US-10-027-632-78585	Sequence 78585, A	C 456	15	1.5	568	15	US-10-027-632-108903	Sequence 108903, A
C 384	15	1.5	481	15	US-10-027-632-78586	Sequence 78586, A	457	15	1.5	569	15	US-10-027-632-241522	Sequence 241522, A
C 385	15	1.5	481	15	US-10-027-632-78587	Sequence 78587, A	458	15	1.5	569	15	US-10-027-632-241523	Sequence 241523, A
C 386	15	1.5	481	15	US-10-027-632-78588	Sequence 78588, A	C 459	15	1.5	569	15	US-10-027-632-270154	Sequence 270154, A
C 387	15	1.5	481	15	US-10-027-632-195074	Sequence 195074, A	C 460	15	1.5	572	15	US-10-198-846-11211	Sequence 11211, A
C 388	15	1.5	481	15	US-10-027-632-195179	Sequence 195179, A	C 461	15	1.5	573	10	US-09-864-761-6605	Sequence 6605, Ap
C 389	15	1.5	481	15	US-10-198-846-5181	Sequence 5181, Ap	C 462	15	1.5	573	15	US-10-027-632-173367	Sequence 173367, A
C 390	15	1.5	486	10	US-09-215-450-2	Sequence 2, Appli	C 463	15	1.5	574	15	US-10-027-632-262947	Sequence 262947, A
C 391	15	1.5	486	15	US-10-027-632-82727	Sequence 82727, A	C 464	15	1.5	574	15	US-10-027-632-262948	Sequence 262948, A
C 392	15	1.5	486	15	US-10-027-632-302305	Sequence 302305, A	C 465	15	1.5	574	15	US-10-027-632-262949	Sequence 262949, A
C 393	15	1.5	492	15	US-10-198-846-2286	Sequence 2286, Ap	C 466	15	1.5	574	15	US-10-027-632-262950	Sequence 262950, A
C 394	15	1.5	493	11	US-09-604-287A-53	Sequence 53, Appl	C 467	15	1.5	574	15	US-10-027-632-262951	Sequence 262951, A
C 395	15	1.5	493	12	US-09-339-338-53	Sequence 53, Appl	C 468	15	1.5	574	15	US-10-027-632-278540	Sequence 278540, A
C 396	15	1.5	493	12	US-09-551-621-53	Sequence 53, Appl	469	15	1.5	575	15	US-10-027-632-303576	Sequence 303576, A
C 397	15	1.5	493	15	US-10-007-805-53	Sequence 53, Appl	470	15	1.5	576	15	US-10-027-632-246419	Sequence 246419, A
C 398	15	1.5	493	15	US-10-027-632-308704	Sequence 308704, A	471	15	1.5	576	15	US-10-027-632-246420	Sequence 246420, A
C 399	15	1.5	493	15	US-10-027-632-308705	Sequence 308705, A	472	15	1.5	576	15	US-10-027-632-246421	Sequence 246421, A
C 400	15	1.5	493	15	US-10-076-622-53	Sequence 53, Appl	473	15	1.5	577	15	US-10-027-632-37309	Sequence 37309, A
C 401	15	1.5	494	11	US-09-833-381-1944	Sequence 1944, Ap	C 474	15	1.5	579	15	US-10-027-632-224998	Sequence 224998, A
C 402	15	1.5	494	15	US-10-027-632-194597	Sequence 194597, A	C 475	15	1.5	579	15	US-10-038-010-19	Sequence 19, Appl
C 403	15	1.5	494	15	US-10-027-632-277532	Sequence 277532, A	476	15	1.5	580	10	US-09-864-761-9860	Sequence 9860, Ap
C 404	15	1.5	494	15	US-10-027-632-277533	Sequence 277533, A	477	15	1.5	580	15	US-10-027-632-40624	Sequence 40624, A
C 405	15	1.5	498	15	US-10-027-632-286705	Sequence 286705, A	C 478	15	1.5	581	15	US-10-027-632-229250	Sequence 229250, A
C 406	15	1.5	498	15	US-10-027-632-323086	Sequence 323086, A	C 479	15	1.5	584	15	US-10-027-632-258956	Sequence 258956, A
C 407	15	1.5	499	11	US-09-974-300-4768	Sequence 4768, Ap	480	15	1.5	585	11	US-09-738-626-2611	Sequence 2611, Ap
C 408	15	1.5	502	15	US-10-027-632-234123	Sequence 234123, A	481	15	1.5	585	15	US-10-027-632-187337	Sequence 187337, A
C 409	15	1.5	503	15	US-10-027-632-324037	Sequence 324037, A	482	15	1.5	585	15	US-10-027-632-187338	Sequence 187338, A
C 410	15	1.5	503	15	US-10-027-632-324038	Sequence 324038, A	C 483	15	1.5	585	15	US-10-027-632-223691	Sequence 223691, A
C 411	15	1.5	505	15	US-10-027-632-43441	Sequence 43441, A	484	15	1.5	586	15	US-10-027-632-209436	Sequence 209436, A
C 412	15	1.5	505	15	US-10-027-632-133808	Sequence 133808, A	C 485	15	1.5	589	15	US-10-062-254-193	Sequence 193, App
C 413	15	1.5	508	15	US-10-027-632-188232	Sequence 188232, A	C 486	15	1.5	589	15	US-10-027-632-269865	Sequence 269865, A
C 414	15	1.5	508	15	US-10-027-632-286282	Sequence 286282, A	C 487	15	1.5	590	15	US-10-027-632-290700	Sequence 290700, A
C 415	15	1.5	510	15	US-10-027-632-281761	Sequence 281761, A	C 488	15	1.5	591	15	US-10-027-632-213604	Sequence 213604, A
C 416	15	1.5	514	10	US-09-864-761-7848	Sequence 7848, Ap	C 489	15	1.5	591	15	US-10-027-632-213605	Sequence 213605, A
C 417	15	1.5	514	15	US-10-027-632-141263	Sequence 141263, A	C 490	15	1.5	591	15	US-10-027-632-213606	Sequence 213606, A
C 418	15	1.5	515	15	US-10-027-632-186407	Sequence 186407, A	491	15	1.5	592	15	US-10-027-632-131125	Sequence 131125, A
C 419	15	1.5	515	15	US-10-027-632-186408	Sequence 186408, A	492	15	1.5	594	10	US-09-864-761-6622	Sequence 6622, Ap
C 420	15	1.5	520	15	US-10-027-632-187838	Sequence 187838, A	C 493	15	1.5	596	15	US-10-027-632-90656	Sequence 90656, A
C 421	15	1.5	520	15	US-10-027-632-187839	Sequence 187839, A	C 494	15	1.5	596	15	US-10-027-632-90657	Sequence 90657, A
C 422	15	1.5	521	15	US-10-027-632-233717	Sequence 233717, A	495	15	1.5	596	15	US-10-027-632-247885	Sequence 247885, A
C 423	15	1.5	523	15	US-10-027-632-288815	Sequence 288815, A	496	15	1.5	598	15	US-10-027-632-190169	Sequence 190169, A
C 424	15	1.5	525	15	US-10-027-632-109621	Sequence 109621, A	C 497	15	1.5	598	15	US-10-027-632-233741	Sequence 233741, A
C 425	15	1.5	533	15	US-10-027-632-293138	Sequence 293138, A	C 498	15	1.5	599	15	US-10-027-632-5861	Sequence 5861, Ap
C 426	15	1.5	533	15	US-10-078-770-105	Sequence 105, App	C 499	15	1.5	599	15	US-10-027-632-5862	Sequence 5862, Ap
C 427	15	1.5	535	11	US-09-879-536-298	Sequence 298, App	C 500	15	1.5	599	15	US-10-027-632-5863	Sequence 5863, Ap
C 428	15	1.5	536	15	US-10-027-632-278539	Sequence 278539, A	501	15	1.5	603	15	US-10-027-632-88521	Sequence 88521, A
C 429	15	1.5	538	15	US-10-027-632-193906	Sequence 193906, A	502	15	1.5	603	15	US-10-027-632-303821	Sequence 303821, A
C 430	15	1.5	538	15	US-10-027-632-193907	Sequence 193907, A	503	15	1.5	604	15	US-10-027-632-17929	Sequence 17929, A
C 431	15	1.5	540	15	US-10-027-632-5237	Sequence 5237, Ap	C 504	15	1.5	604	15	US-10-027-632-136044	Sequence 136044, A
C 432	15	1.5	542	11	US-09-920-300A-1402	Sequence 1402, Ap	C 505	15	1.5	605	15	US-10-027-632-34757	Sequence 34757, A
C 433	15	1.5	542	15	US-10-033-528-1402	Sequence 1402, Ap	506	15	1.5	606	15	US-10-027-632-5882	Sequence 5882, Ap
C 434	15	1.5	542	15	US-10-027-632-116203	Sequence 116203, A	C 507	15	1.5	606	15	US-10-027-632-224689	Sequence 224689, A
C 435	15	1.5	543	15	US-10-027-632-89398	Sequence 89398, A	C 508	15	1.5	607	15	US-10-027-632-184953	Sequence 184953, A
C 436	15	1.5	543	15	US-10-027-632-303993	Sequence 303993, A	C 509	15	1.5	607	15	US-10-027-632-263628	Sequence 263628, A
C 437	15	1.5	545	15	US-10-027-632-57461	Sequence 57461, A	C 510	15	1.5	607	15	US-10-027-632-263629	Sequence 263629, A
C 438	15	1.5	546	15	US-10-027-632-70604	Sequence 70604, A	511	15	1.5	607	15	US-10-027-632-279422	Sequence 279422, A
C 439	15	1.5	546	15	US-10-027-632-179375	Sequence 179375, A	512	15	1.5	607	15	US-10-027-632-279423	Sequence 279423, A
C 440	15	1.5	546	15	US-10-027-632-200029	Sequence 200029, A	513	15	1.5	609	15	US-10-027-632-113221	Sequence 113221, A
C 441	15	1.5	546	15	US-10-171-581-131	Sequence 131, App	514	15	1.5	609	15	US-10-027-632-242009	Sequence 242009, A
C 442	15	1.5	548	10	US-09-864-761-11871	Sequence 11871, A	515	15	1.5	609	15	US-10-027-632-242010	Sequence 242010, A
C 443	15	1.5	551	15	US-10-027-632-140043	Sequence 140043, A	516	15	1.5	609	15	US-10-027-632-242011	Sequence 242011, A
C 444	15	1.5	551	15	US-10-027-632-91318	Sequence 91318, A	517	15	1.5	609	15	US-10-027-632-259426	Sequence 259426, A
C 445	15	1.5	551	15	US-10-027-632-278384	Sequence 278384, A	C 518	15	1.5	610	15	US-10-027-632-237688	Sequence 237688, A
C 446	15	1.5	551	15	US-10-027-632-278385	Sequence 278385, A	C 519	15	1.5	611	15	US-10-027-632-179281	Sequence 179281, A
C 447	15	1.5	551	15	US-10-027-632-322285	Sequence 322285, A	C 520	15	1.5	611	15	US-10-027-632-179282	Sequence 179282, A
C 448	15	1.5	556	11	US-09-796-692-8604	Sequence 8604, Ap	C 521	15	1.5	611	15	US-10-027-632-179283	Sequence 179283, A
C 449	15	1.5	556	15	US-10-040-862-8604	Sequence 8604, Ap	C 522	15	1.5	611	15	US-10-027-632-223058	Sequence 223058, A
C 450	15	1.5	559	15	US-10-027-632-180569	Sequence 180569, A	C 523	15	1.5	611	15	US-10-027-632-223059	Sequence 223059, A
C 451	15	1.5	559	15	US-10-027-632-223111	Sequence 223111, A	C 524	15	1.5	611	15	US-10-027-632-253522	Sequence 253522, A
C 452	15	1.5	559	15	US-10-027-632-247178	Sequence 247178, A	C 525	15	1.5	611	15	US-10-027-632-269792	Sequence 269792, A
C 453	15	1.5	567	15	US-10-027-632-84622	Sequence 84622, A	526	15	1.5	614	15	US-10-027-632-251058	Sequence 251058, A
C 454	15	1.5	567	15	US-10-027-632-272018	Sequence 272018, A	527	15	1.5	614	15	US-10-027-632-251059	Sequence 251059, A



528	15	1.5	614	15	US-10-027-632-251060	Sequence 251060,	601	15	1.5	673	11	US-09-070-927A-704	Sequence 704, App
c 529	15	1.5	616	15	US-10-027-632-44973	Sequence 44973, A	602	15	1.5	675	15	US-10-027-632-213628	Sequence 213628,
c 530	15	1.5	616	15	US-10-027-632-131416	Sequence 131416,	c 603	15	1.5	676	15	US-10-027-632-218340	Sequence 218340,
c 531	15	1.5	616	15	US-10-027-632-131417	Sequence 131417,	604	15	1.5	676	15	US-10-027-632-247484	Sequence 247484,
c 532	15	1.5	616	15	US-10-027-632-131418	Sequence 131418,	605	15	1.5	679	15	US-10-027-632-139120	Sequence 139120,
c 533	15	1.5	616	15	US-10-027-632-131419	Sequence 131419,	606	15	1.5	681	11	US-09-989-920-53	Sequence 53, Appl
c 534	15	1.5	616	15	US-10-027-632-217396	Sequence 217396,	607	15	1.5	681	15	US-10-027-632-184945	Sequence 184945,
c 535	15	1.5	616	15	US-10-027-632-217397	Sequence 217397,	c 608	15	1.5	682	15	US-10-027-632-229908	Sequence 229908,
c 536	15	1.5	616	15	US-10-027-632-217397	Sequence 217397,	c 609	15	1.5	682	15	US-10-027-632-229909	Sequence 229909,
c 537	15	1.5	617	15	US-10-027-632-250750	Sequence 250750,	c 610	15	1.5	684	15	US-10-027-632-14605	Sequence 14605, A
c 538	15	1.5	617	15	US-10-027-632-250751	Sequence 250751,	c 611	15	1.5	684	15	US-10-027-632-14606	Sequence 14606, A
c 539	15	1.5	617	15	US-10-027-632-271201	Sequence 271201,	612	15	1.5	689	15	US-10-027-632-209264	Sequence 209264,
c 540	15	1.5	619	15	US-10-027-632-175056	Sequence 175056,	613	15	1.5	693	11	US-09-738-626-2516	Sequence 2516, Ap
541	15	1.5	622	15	US-10-027-632-55050	Sequence 55050, A	614	15	1.5	695	10	US-09-828-644-52	Sequence 52, Appl
542	15	1.5	622	15	US-10-027-632-61935	Sequence 61935, A	615	15	1.5	703	15	US-10-027-632-107098	Sequence 107098,
543	15	1.5	622	15	US-10-027-632-61935	Sequence 61935, A	616	15	1.5	703	15	US-10-027-632-126980	Sequence 126980,
544	15	1.5	622	15	US-10-027-632-182670	Sequence 182670,	617	15	1.5	703	15	US-10-027-632-126981	Sequence 126981,
545	15	1.5	622	15	US-10-027-632-182671	Sequence 182671,	618	15	1.5	703	15	US-10-027-632-137321	Sequence 137321,
546	15	1.5	622	15	US-10-027-632-182672	Sequence 182672,	c 619	15	1.5	703	15	US-10-027-632-206082	Sequence 206082,
547	15	1.5	622	15	US-10-027-632-182673	Sequence 182673,	c 620	15	1.5	706	15	US-10-027-632-116529	Sequence 116529,
548	15	1.5	622	15	US-10-027-632-300435	Sequence 300435,	c 621	15	1.5	707	15	US-10-027-632-127727	Sequence 127727,
c 549	15	1.5	624	15	US-10-027-632-300436	Sequence 300436,	c 622	15	1.5	713	15	US-10-027-632-110822	Sequence 110822,
c 550	15	1.5	624	15	US-10-027-632-220559	Sequence 220559,	c 623	15	1.5	715	15	US-10-027-632-137711	Sequence 137711,
c 551	15	1.5	625	15	US-10-027-632-295780	Sequence 295780,	c 624	15	1.5	715	15	US-10-027-632-165596	Sequence 165596,
c 552	15	1.5	625	15	US-10-027-632-68516	Sequence 68516, A	625	15	1.5	717	15	US-10-027-632-26405	Sequence 26405, A
c 553	15	1.5	625	15	US-10-027-632-68517	Sequence 68517, A	c 626	15	1.5	720	10	US-09-910-943-105	Sequence 105, App
c 554	15	1.5	625	15	US-10-027-632-192335	Sequence 192335,	c 627	15	1.5	725	15	US-10-027-632-157268	Sequence 157268,
c 555	15	1.5	625	15	US-10-027-632-192336	Sequence 192336,	628	15	1.5	725	15	US-10-027-632-163230	Sequence 163230,
c 556	15	1.5	625	15	US-10-027-632-192337	Sequence 192337,	c 629	15	1.5	729	10	US-09-910-943-104	Sequence 104, App
c 557	15	1.5	627	15	US-10-027-632-6039	Sequence 6039, Ap	630	15	1.5	729	15	US-10-027-632-128732	Sequence 128732,
558	15	1.5	627	15	US-10-027-632-64174	Sequence 64174, A	631	15	1.5	732	11	US-09-995-598-33	Sequence 33, Appl
559	15	1.5	627	15	US-10-027-632-64975	Sequence 64975, A	c 632	15	1.5	732	15	US-10-027-632-128731	Sequence 128731,
560	15	1.5	627	15	US-10-027-632-107874	Sequence 107874,	633	15	1.5	734	15	US-10-027-632-12888	Sequence 12888, A
561	15	1.5	628	15	US-10-027-632-238847	Sequence 238847,	c 634	15	1.5	740	15	US-10-027-632-99718	Sequence 99718, A
562	15	1.5	629	15	US-10-027-632-220781	Sequence 220781,	c 635	15	1.5	745	15	US-10-027-632-21881	Sequence 21881, A
c 563	15	1.5	630	15	US-10-027-632-223219	Sequence 223219,	c 636	15	1.5	748	10	US-09-910-943-172	Sequence 172, App
c 564	15	1.5	631	15	US-10-027-632-46216	Sequence 46216, A	c 637	15	1.5	749	15	US-10-027-632-137322	Sequence 137322,
c 565	15	1.5	631	15	US-10-027-632-46217	Sequence 46217, A	c 638	15	1.5	749	15	US-10-027-632-137323	Sequence 137323,
c 566	15	1.5	631	15	US-10-027-632-182370	Sequence 182370,	c 639	15	1.5	757	15	US-10-027-632-153307	Sequence 153307,
c 567	15	1.5	631	15	US-10-027-632-309007	Sequence 309007,	c 640	15	1.5	757	15	US-10-027-632-153308	Sequence 153308,
568	15	1.5	633	15	US-10-027-632-113626	Sequence 113626,	641	15	1.5	758	10	US-09-910-943-55	Sequence 55, Appl
c 569	15	1.5	635	15	US-10-027-632-143146	Sequence 143146,	c 642	15	1.5	761	15	US-10-062-254-195	Sequence 195, App
c 570	15	1.5	635	15	US-10-027-632-281946	Sequence 281946,	643	15	1.5	762	11	US-09-995-598-39	Sequence 39, Appl
571	15	1.5	635	15	US-10-198-846-8146	Sequence 8146, Ap	c 644	15	1.5	766	10	US-09-864-761-17873	Sequence 17873, A
572	15	1.5	636	11	US-09-738-626-2734	Sequence 2734, Ap	645	15	1.5	770	11	US-09-974-300-2057	Sequence 2057, Ap
573	15	1.5	637	15	US-10-027-632-225670	Sequence 225670,	646	15	1.5	778	15	US-10-144-929-58	Sequence 58, Appl
574	15	1.5	637	15	US-10-027-632-293496	Sequence 293496,	c 647	15	1.5	781	15	US-10-027-632-987	Sequence 987, App
575	15	1.5	638	15	US-10-027-632-114302	Sequence 114302,	c 648	15	1.5	789	11	US-09-903-410-25	Sequence 25, Appl
576	15	1.5	638	15	US-10-027-632-137669	Sequence 137669,	c 649	15	1.5	789	15	US-10-027-805-25	Sequence 25, Appl
577	15	1.5	641	15	US-10-027-632-46858	Sequence 46858, A	c 650	15	1.5	789	15	US-10-027-804-25	Sequence 25, Appl
578	15	1.5	644	15	US-10-027-632-104913	Sequence 104913,	651	15	1.5	795	10	US-09-815-242-7524	Sequence 7524, Ap
579	15	1.5	644	15	US-10-027-632-133338	Sequence 133338,	652	15	1.5	801	15	US-10-027-632-153514	Sequence 153514,
c 580	15	1.5	647	15	US-10-027-632-80247	Sequence 80247, A	c 653	15	1.5	804	11	US-09-966-881-56	Sequence 56, Appl
581	15	1.5	647	15	US-10-027-632-211725	Sequence 211725,	c 654	15	1.5	810	11	US-09-966-881-23	Sequence 23, Appl
582	15	1.5	648	10	US-09-867-550-765	Sequence 765, App	c 655	15	1.5	810	15	US-10-198-846-12699	Sequence 12699, A
583	15	1.5	650	15	US-10-027-632-213333	Sequence 213333,	c 656	15	1.5	815	15	US-10-027-632-168064	Sequence 168064,
c 584	15	1.5	650	15	US-10-027-632-233010	Sequence 233010,	657	15	1.5	817	15	US-10-027-632-31247	Sequence 31247, A
585	15	1.5	650	15	US-10-027-632-300691	Sequence 300691,	c 658	15	1.5	817	15	US-10-027-632-31248	Sequence 31248, A
586	15	1.5	652	11	US-09-864-864-156	Sequence 156, App	c 659	15	1.5	822	15	US-10-027-632-807	Sequence 807, App
c 587	15	1.5	653	15	US-10-027-632-79820	Sequence 79820, A	660	15	1.5	822	15	US-10-027-632-170606	Sequence 170606,
c 588	15	1.5	653	15	US-10-027-632-80742	Sequence 80742, A	661	15	1.5	824	15	US-10-027-632-167201	Sequence 167201,
c 589	15	1.5	653	15	US-10-027-632-80871	Sequence 80871, A	662	15	1.5	824	15	US-10-027-632-167202	Sequence 167202,
c 590	15	1.5	655	15	US-10-071-766-62	Sequence 62, Appl	663	15	1.5	825	15	US-10-027-632-154899	Sequence 154899,
c 591	15	1.5	655	15	US-10-027-632-227367	Sequence 227367,	c 664	15	1.5	826	15	US-10-027-632-127913	Sequence 127913,
592	15	1.5	662	15	US-10-027-632-277062	Sequence 277062,	c 665	15	1.5	830	15	US-10-027-632-171402	Sequence 171402,
593	15	1.5	663	12	US-09-899-495-46	Sequence 46, Appl	c 666	15	1.5	830	15	US-10-027-632-171403	Sequence 171403,
c 594	15	1.5	663	15	US-10-027-632-250796	Sequence 250796,	667	15	1.5	833	15	US-10-198-846-9335	Sequence 9335, Ap
c 595	15	1.5	663	15	US-10-027-632-250797	Sequence 250797,	668	15	1.5	836	15	US-10-027-632-165942	Sequence 165942,
596	15	1.5	665	15	US-10-027-632-276047	Sequence 276047,	669	15	1.5	836	15	US-10-027-632-165943	Sequence 165943,
597	15	1.5	666	15	US-10-027-632-120689	Sequence 120689,	670	15	1.5	836	15	US-10-027-632-165944	Sequence 165944,
598	15	1.5	669	15	US-10-027-632-219605	Sequence 219605,	671	15	1.5	839	15	US-10-198-846-5090	Sequence 5090, Ap
c 599	15	1.5	669	15	US-10-259-430-11	Sequence 11, Appl	672	15	1.5	840	15	US-10-027-632-166144	Sequence 166144,
c 600	15	1.5	672	15	US-10-027-632-101602	Sequence 101602,	673	15	1.5	840	15	US-10-027-632-166145	Sequence 166145,



c 674	15	1.5	842	15	US-10-027-632-146868	Sequence 146868,	c 747	15	1.5	1375	15	US-10-091-572-907	Sequence 907, App
c 675	15	1.5	843	15	US-10-027-632-159952	Sequence 159952,	c 748	15	1.5	1375	15	US-10-091-572-908	Sequence 908, App
676	15	1.5	851	15	US-10-027-632-8924	Sequence 8924, Ap	749	15	1.5	1383	15	US-10-033-078-3	Sequence 3, Appli
677	15	1.5	855	15	US-10-027-632-9798	Sequence 9798, Ap	750	15	1.5	1393	11	US-09-974-300-1687	Sequence 1687, Ap
678	15	1.5	855	15	US-10-027-632-163615	Sequence 163615,	c 751	15	1.5	1399	11	US-09-974-300-1163	Sequence 1163, Ap
c 679	15	1.5	861	15	US-10-027-632-173997	Sequence 173997,	c 752	15	1.5	1407	11	US-09-974-300-2053	Sequence 2053, Ap
c 680	15	1.5	861	15	US-10-027-632-173998	Sequence 173998,	753	15	1.5	1446	11	US-09-938-842A-802	Sequence 802, App
c 681	15	1.5	871	15	US-10-027-632-171373	Sequence 171373,	754	15	1.5	1476	15	US-10-322-149-13	Sequence 13, Appl
c 682	15	1.5	871	15	US-10-027-632-171374	Sequence 171374,	755	15	1.5	1492	15	US-10-078-770-113	Sequence 113, App
683	15	1.5	900	11	US-09-738-626-643	Sequence 643, App	c 756	15	1.5	1507	8	US-08-781-986A-361	Sequence 361, App
684	15	1.5	907	15	US-10-004-717-65	Sequence 65, Appl	c 757	15	1.5	1542	11	US-09-860-670-66	Sequence 66, Appl
c 685	15	1.5	908	15	US-10-027-632-159048	Sequence 159048,	c 758	15	1.5	1549	15	US-10-001-835-102	Sequence 102, App
c 686	15	1.5	908	15	US-10-144-929-54	Sequence 54, Appl	759	15	1.5	1599	11	US-09-938-842A-1762	Sequence 1762, Ap
687	15	1.5	917	15	US-10-198-846-11251	Sequence 11251, A	c 760	15	1.5	1603	15	US-10-027-632-260006	Sequence 260006,
c 688	15	1.5	920	15	US-10-198-846-1396	Sequence 3396, Ap	761	15	1.5	1605	12	US-09-986-480-21	Sequence 21, Appl
c 689	15	1.5	928	15	US-10-027-632-33634	Sequence 33634, A	762	15	1.5	1605	12	US-09-895-298-42	Sequence 42, Appl
690	15	1.5	942	11	US-09-738-626-3210	Sequence 3210, Ap	763	15	1.5	1606	15	US-10-098-841-323	Sequence 323, App
691	15	1.5	944	15	US-10-027-632-159953	Sequence 159953,	764	15	1.5	1614	11	US-09-917-800A-1525	Sequence 1525, Ap
692	15	1.5	982	15	US-10-027-632-265049	Sequence 265049,	c 765	15	1.5	1629	15	US-10-027-632-262689	Sequence 262689,
c 693	15	1.5	992	11	US-09-974-300-1716	Sequence 1716, Ap	c 766	15	1.5	1744	11	US-09-833-381-1943	Sequence 1943, Ap
c 694	15	1.5	996	10	US-09-900-715-1	Sequence 1, Appli	767	15	1.5	1750	10	US-09-925-301-403	Sequence 403, App
695	15	1.5	998	15	US-10-106-698-1599	Sequence 1599, Ap	c 768	15	1.5	1797	10	US-09-875-811-11	Sequence 11, Appl
696	15	1.5	1011	11	US-09-938-842A-1719	Sequence 1719, Ap	c 769	15	1.5	1818	10	US-09-875-811-7	Sequence 7, Appli
697	15	1.5	1021	11	US-09-954-456-186	Sequence 186, App	c 770	15	1.5	1823	15	US-10-171-581-126	Sequence 126, App
698	15	1.5	1021	11	US-09-954-456-941	Sequence 941, App	c 771	15	1.5	1824	15	US-10-044-090-372	Sequence 372, App
699	15	1.5	1050	10	US-09-886-468-4	Sequence 4, Appli	c 772	15	1.5	1840	11	US-09-895-913A-139	Sequence 139, App
700	15	1.5	1078	10	US-09-923-779-147	Sequence 147, App	c 773	15	1.5	1849	15	US-10-027-632-39099	Sequence 39099, A
c 701	15	1.5	1078	11	US-09-835-948-28	Sequence 28, Appl	c 774	15	1.5	1849	15	US-10-027-632-39100	Sequence 39100, A
702	15	1.5	1078	11	US-09-835-948-29	Sequence 29, Appl	c 775	15	1.5	1849	15	US-10-027-632-39101	Sequence 39101, A
703	15	1.5	1078	15	US-10-097-340-190	Sequence 190, App	c 776	15	1.5	1849	15	US-10-027-632-39102	Sequence 39102, A
704	15	1.5	1078	15	US-10-171-311-140	Sequence 140, App	c 777	15	1.5	1859	11	US-09-915-582-23	Sequence 23, Appl
705	15	1.5	1078	15	US-10-205-823-267	Sequence 267, App	c 778	15	1.5	1866	10	US-09-875-811-3	Sequence 3, Appli
706	15	1.5	1078	15	US-10-177-293-299	Sequence 299, App	c 779	15	1.5	1868	11	US-09-938-842A-5068	Sequence 5068, Ap
707	15	1.5	1094	11	US-09-070-927A-640	Sequence 640, App	c 780	15	1.5	1936	11	US-09-822-830A-291	Sequence 291, App
708	15	1.5	1096	11	US-09-974-300-2393	Sequence 2393, Ap	c 781	15	1.5	1937	11	US-09-822-830A-364	Sequence 364, App
709	15	1.5	1107	15	US-10-156-761-3873	Sequence 3873, Ap	c 782	15	1.5	1946	15	US-10-006-852-9	Sequence 9, Appli
710	15	1.5	1110	15	US-10-027-632-118261	Sequence 118261,	783	15	1.5	1967	15	US-10-027-632-264119	Sequence 264119,
c 711	15	1.5	1113	11	US-09-993-844-9	Sequence 9, Appli	c 784	15	1.5	1970	12	US-09-952-680A-11	Sequence 11, Appl
712	15	1.5	1125	11	US-09-798-584-2	Sequence 2, Appli	c 785	15	1.5	1971	10	US-09-875-811-9	Sequence 9, Appli
713	15	1.5	1127	15	US-10-171-311-138	Sequence 138, App	c 786	15	1.5	1971	10	US-09-942-446-3	Sequence 3, Appli
714	15	1.5	1127	15	US-10-205-823-265	Sequence 265, App	787	15	1.5	1973	15	US-10-027-632-98458	Sequence 98458, A
715	15	1.5	1127	15	US-10-102-524-1705	Sequence 1705, Ap	788	15	1.5	1973	15	US-10-027-632-98459	Sequence 98459, A
c 716	15	1.5	1137	11	US-09-974-300-2309	Sequence 2309, Ap	789	15	1.5	1973	15	US-10-027-632-98460	Sequence 98460, A
c 717	15	1.5	1144	11	US-09-974-300-804	Sequence 804, App	790	15	1.5	1977	11	US-09-938-842A-1120	Sequence 1120, Ap
718	15	1.5	1167	11	US-09-738-626-3042	Sequence 3042, Ap	791	15	1.5	1982	15	US-10-027-632-97898	Sequence 97898, A
719	15	1.5	1173	15	US-10-027-632-116754	Sequence 116754,	792	15	1.5	1982	15	US-10-027-632-97899	Sequence 97899, A
720	15	1.5	1173	15	US-10-106-698-2123	Sequence 2123, Ap	c 793	15	1.5	1992	10	US-09-875-811-5	Sequence 5, Appli
c 721	15	1.5	1174	11	US-09-938-842A-3677	Sequence 3677, Ap	794	15	1.5	1992	15	US-10-027-632-263202	Sequence 263202,
c 722	15	1.5	1176	15	US-10-027-632-119084	Sequence 119084,	795	15	1.5	1992	15	US-10-027-632-263203	Sequence 263203,
723	15	1.5	1186	15	US-10-027-632-123952	Sequence 123952,	796	15	1.5	1992	15	US-10-027-632-263204	Sequence 263204,
c 724	15	1.5	1224	15	US-10-027-632-213607	Sequence 213607,	c 797	15	1.5	2000	11	US-09-938-842A-3396	Sequence 3396, Ap
c 725	15	1.5	1239	9	US-09-935-061-5	Sequence 5, Appli	798	15	1.5	2000	11	US-09-938-842A-3712	Sequence 3712, Ap
c 726	15	1.5	1239	9	US-09-935-061-7	Sequence 7, Appli	799	15	1.5	2000	11	US-09-938-842A-5097	Sequence 5097, Ap
c 728	15	1.5	1244	15	US-10-027-632-122576	Sequence 122576,	800	15	1.5	2003	11	US-09-887-576-306	Sequence 306, App
729	15	1.5	1251	9	US-09-935-061-9	Sequence 9, Appli	801	15	1.5	2004	11	US-09-887-576-253	Sequence 253, App
730	15	1.5	1261	15	US-10-027-632-124450	Sequence 124450,	c 802	15	1.5	2004	11	US-09-764-877-4015	Sequence 4015, Ap
731	15	1.5	1265	15	US-10-027-632-254386	Sequence 254386,	c 803	15	1.5	2004	11	US-09-764-877-4016	Sequence 4016, Ap
732	15	1.5	1267	10	US-09-925-301-188	Sequence 188, App	c 804	15	1.5	2015	15	US-10-225-567A-51	Sequence 51, Appl
733	15	1.5	1277	15	US-10-027-632-250712	Sequence 250712,	c 805	15	1.5	2026	15	US-10-027-632-97254	Sequence 97254, A
734	15	1.5	1288	15	US-10-027-632-123418	Sequence 123418,	806	15	1.5	2026	15	US-10-027-632-97255	Sequence 97255, A
c 734	15	1.5	1301	15	US-10-027-632-177476	Sequence 177476,	c 807	15	1.5	2040	10	US-09-875-811-1	Sequence 1, Appli
735	15	1.5	1302	11	US-09-974-300-2641	Sequence 2641, Ap	808	15	1.5	2040	15	US-10-027-632-98287	Sequence 98287, A
736	15	1.5	1332	15	US-10-198-846-13192	Sequence 13192, A	809	15	1.5	2065	15	US-10-027-632-98139	Sequence 98139, A
c 737	15	1.5	1333	15	US-10-027-632-124062	Sequence 124062,	c 810	15	1.5	2115	10	US-09-815-242-4279	Sequence 4279, Ap
c 738	15	1.5	1333	15	US-10-027-632-124063	Sequence 124063,	811	15	1.5	2119	10	US-09-764-869-2439	Sequence 2439, Ap
739	15	1.5	1343	15	US-10-027-632-205444	Sequence 205444,	812	15	1.5	2119	15	US-10-091-504-2439	Sequence 2439, Ap
c 740	15	1.5	1346	11	US-09-828-302-9	Sequence 9, Appli	c 813	15	1.5	2121	11	US-09-738-626-1463	Sequence 1463, Ap
741	15	1.5	1358	11	US-09-822-830A-212	Sequence 212, App	c 814	15	1.5	2122	12	US-09-955-999-59	Sequence 59, Appl
742	15	1.5	1371	15	US-10-156-761-6160	Sequence 6160, Ap	815	15	1.5	2142	10	US-09-764-869-2440	Sequence 2440, Ap
c 743	15	1.5	1374	12	US-09-764-891-10088	Sequence 10088, A	c 816	15	1.5	2142	10	US-09-815-242-8224	Sequence 8224, Ap
c 744	15	1.5	1374	15	US-10-091-572-906	Sequence 906, App	817	15	1.5	2142	15	US-10-091-504-2440	Sequence 2440, Ap
c 745	15	1.5	1375	12	US-09-764-891-10089	Sequence 10089, A	818	15	1.5	2146	15	US-10-027-632-97126	Sequence 97126, A
c 746	15	1.5	1375	12	US-09-764-891-10090	Sequence 10090, A	819	15	1.5	2146	15	US-10-027-632-97127	Sequence 97127, A

C 820	15	1.5	2158	15	US-10-027-632-262660	Sequence 262660,	C 893	15	1.5	3501	15	US-10-140-474-37	Sequence 37, Appl
C 821	15	1.5	2158	15	US-10-027-632-262661	Sequence 262661,	C 894	15	1.5	3501	15	US-10-142-431-37	Sequence 37, Appl
C 822	15	1.5	2158	15	US-10-027-632-262662	Sequence 262662,	C 895	15	1.5	3501	15	US-10-143-114-37	Sequence 37, Appl
823	15	1.5	2162	15	US-10-027-632-254428	Sequence 254428,	C 896	15	1.5	3501	15	US-10-140-002-37	Sequence 37, Appl
C 824	15	1.5	2179	15	US-10-071-766-63	Sequence 63, Appl	C 897	15	1.5	3501	15	US-10-142-419-37	Sequence 37, Appl
C 825	15	1.5	2268	15	US-10-037-270-909	Sequence 909, App	C 898	15	1.5	3501	15	US-10-123-262-37	Sequence 37, Appl
C 826	15	1.5	2299	15	US-10-198-846-13500	Sequence 13500, A	C 899	15	1.5	3501	15	US-10-142-423-37	Sequence 37, Appl
C 827	15	1.5	2305	15	US-10-309-515-23	Sequence 23, Appl	C 900	15	1.5	3501	15	US-10-121-050-37	Sequence 37, Appl
828	15	1.5	2312	11	US-09-798-584-5	Sequence 5, Appl	C 901	15	1.5	3501	15	US-10-141-755-37	Sequence 37, Appl
C 829	15	1.5	2340	11	US-09-738-626-1206	Sequence 1206, Ap	C 902	15	1.5	3501	15	US-10-143-032-37	Sequence 37, Appl
C 830	15	1.5	2345	15	US-10-222-009-12	Sequence 12, Appl	C 903	15	1.5	3501	15	US-10-123-108-37	Sequence 37, Appl
C 831	15	1.5	2357	12	US-09-809-391-262	Sequence 262, App	C 904	15	1.5	3501	15	US-10-123-236-37	Sequence 37, Appl
C 832	15	1.5	2367	11	US-09-995-587A-2	Sequence 2, Appl	C 905	15	1.5	3501	15	US-10-123-261-37	Sequence 37, Appl
C 833	15	1.5	2376	11	US-09-880-107-2227	Sequence 2227, Ap	C 906	15	1.5	3501	15	US-10-140-921-37	Sequence 37, Appl
834	15	1.5	2379	11	US-09-974-300-671	Sequence 671, App	C 907	15	1.5	3501	15	US-10-140-928-37	Sequence 37, Appl
C 835	15	1.5	2379	15	US-10-156-761-1262	Sequence 1262, Ap	C 908	15	1.5	3501	15	US-10-121-045-37	Sequence 37, Appl
C 836	15	1.5	2394	11	US-09-995-587A-3	Sequence 3, Appl	C 909	15	1.5	3501	15	US-10-123-292-37	Sequence 37, Appl
C 837	15	1.5	2404	14	US-10-323-001-1	Sequence 1, Appl	C 910	15	1.5	3501	15	US-10-123-903-37	Sequence 37, Appl
C 838	15	1.5	2404	15	US-10-242-332-1	Sequence 1, Appl	C 911	15	1.5	3501	15	US-10-124-819-37	Sequence 37, Appl
839	15	1.5	2409	11	US-09-938-842A-2259	Sequence 2259, Ap	C 912	15	1.5	3501	15	US-10-124-822-37	Sequence 37, Appl
C 840	15	1.5	2420	10	US-09-875-811-13	Sequence 13, Appl	C 913	15	1.5	3501	15	US-10-140-925-37	Sequence 37, Appl
C 841	15	1.5	2423	15	US-10-198-846-11283	Sequence 11283, A	C 914	15	1.5	3501	15	US-10-160-498-37	Sequence 37, Appl
842	15	1.5	2456	10	US-09-764-898-39	Sequence 39, Appl	C 915	15	1.5	3501	15	US-10-124-824-37	Sequence 37, Appl
C 843	15	1.5	2469	15	US-10-027-632-102088	Sequence 102088,	C 916	15	1.5	3501	15	US-10-127-825A-37	Sequence 37, Appl
C 844	15	1.5	2469	15	US-10-027-632-102089	Sequence 102089,	C 917	15	1.5	3501	15	US-10-127-829A-37	Sequence 37, Appl
C 845	15	1.5	2469	15	US-10-027-632-102090	Sequence 102090,	C 918	15	1.5	3501	15	US-10-127-835A-37	Sequence 37, Appl
846	15	1.5	2470	12	US-09-764-891-9974	Sequence 9974, Ap	C 919	15	1.5	3501	15	US-10-127-839A-37	Sequence 37, Appl
847	15	1.5	2494	11	US-09-925-300-206	Sequence 206, App	C 920	15	1.5	3501	15	US-10-127-901A-37	Sequence 37, Appl
C 848	15	1.5	2514	12	US-09-809-391-261	Sequence 261, App	C 921	15	1.5	3501	15	US-10-128-693A-37	Sequence 37, Appl
C 849	15	1.5	2592	11	US-09-995-587A-4	Sequence 4, Appl	C 922	15	1.5	3501	15	US-10-131-813A-37	Sequence 37, Appl
850	15	1.5	2620	10	US-09-925-302-205	Sequence 205, App	C 923	15	1.5	3501	15	US-10-131-818A-37	Sequence 37, Appl
851	15	1.5	2663	10	US-09-892-867-2	Sequence 2, Appl	C 924	15	1.5	3501	15	US-10-131-823A-37	Sequence 37, Appl
C 852	15	1.5	2735	15	US-10-225-567A-135	Sequence 135, App	C 925	15	1.5	3501	15	US-10-131-824A-37	Sequence 37, Appl
C 853	15	1.5	2886	15	US-10-222-009-11	Sequence 11, Appl	C 926	15	1.5	3501	15	US-10-131-830A-37	Sequence 37, Appl
C 854	15	1.5	2887	11	US-09-964-899-34	Sequence 34, Appl	C 927	15	1.5	3501	15	US-10-131-837A-37	Sequence 37, Appl
855	15	1.5	2902	10	US-09-764-898-19	Sequence 19, Appl	C 928	15	1.5	3501	15	US-10-137-872A-37	Sequence 37, Appl
C 856	15	1.5	2971	15	US-10-002-600-42	Sequence 42, Appl	C 929	15	1.5	3501	15	US-10-147-500-37	Sequence 37, Appl
C 857	15	1.5	2977	12	US-09-919-039-65	Sequence 65, Appl	C 930	15	1.5	3501	15	US-10-147-502-37	Sequence 37, Appl
858	15	1.5	2981	15	US-10-198-846-13446	Sequence 13446, A	C 931	15	1.5	3501	15	US-10-147-515-37	Sequence 37, Appl
859	15	1.5	3054	11	US-09-764-868-64	Sequence 64, Appl	C 932	15	1.5	3501	15	US-10-147-517-37	Sequence 37, Appl
C 860	15	1.5	3111	15	US-10-027-632-114920	Sequence 114920,	C 933	15	1.5	3501	15	US-10-147-526-37	Sequence 37, Appl
C 861	15	1.5	3120	15	US-10-112-286-1	Sequence 1, Appl	C 934	15	1.5	3501	15	US-10-147-527-37	Sequence 37, Appl
862	15	1.5	3124	10	US-09-925-302-340	Sequence 340, App	C 935	15	1.5	3501	15	US-10-121-041-37	Sequence 37, Appl
863	15	1.5	3130	15	US-10-198-846-10416	Sequence 10416, A	C 936	15	1.5	3501	15	US-10-121-043-37	Sequence 37, Appl
C 864	15	1.5	3149	15	US-10-027-632-112984	Sequence 112984,	C 937	15	1.5	3501	15	US-10-121-047-37	Sequence 37, Appl
865	15	1.5	3182	15	US-10-027-632-177173	Sequence 177173,	C 938	15	1.5	3501	15	US-10-123-215-37	Sequence 37, Appl
C 866	15	1.5	3191	11	US-09-989-920-54	Sequence 54, Appl	C 939	15	1.5	3501	15	US-10-123-902-37	Sequence 37, Appl
C 867	15	1.5	3198	15	US-10-027-632-115313	Sequence 115313,	C 940	15	1.5	3501	15	US-10-123-908-37	Sequence 37, Appl
C 868	15	1.5	3201	15	US-10-072-621-5	Sequence 5, Appl	C 941	15	1.5	3501	15	US-10-123-909-37	Sequence 37, Appl
869	15	1.5	3230	15	US-10-027-632-112811	Sequence 112811,	C 942	15	1.5	3501	15	US-10-123-910-37	Sequence 37, Appl
870	15	1.5	3230	15	US-10-027-632-112812	Sequence 112812,	C 943	15	1.5	3501	15	US-10-124-813-37	Sequence 37, Appl
C 871	15	1.5	3358	15	US-10-027-632-112808	Sequence 112808,	C 944	15	1.5	3501	15	US-10-124-817-37	Sequence 37, Appl
C 872	15	1.5	3358	15	US-10-027-632-112809	Sequence 112809,	C 945	15	1.5	3501	15	US-10-125-922-37	Sequence 37, Appl
C 873	15	1.5	3451	10	US-09-811-286-1	Sequence 1, Appl	C 946	15	1.5	3501	15	US-10-125-924-37	Sequence 37, Appl
C 874	15	1.5	3451	11	US-09-954-456-214	Sequence 214, App	C 947	15	1.5	3501	15	US-10-140-860-37	Sequence 37, Appl
C 875	15	1.5	3451	11	US-09-954-456-1598	Sequence 1598, Ap	C 948	15	1.5	3501	15	US-10-142-417-37	Sequence 37, Appl
C 876	15	1.5	3451	11	US-09-990-596-2	Sequence 2, Appl	C 949	15	1.5	3501	15	US-10-147-519-37	Sequence 37, Appl
C 877	15	1.5	3451	12	US-09-811-285-1	Sequence 1, Appl	C 950	15	1.5	3501	15	US-10-157-782-37	Sequence 37, Appl
C 878	15	1.5	3487	15	US-10-114-170-164	Sequence 164, App	C 951	15	1.5	3501	15	US-10-152-395-37	Sequence 37, Appl
C 879	15	1.5	3501	14	US-10-137-866-37	Sequence 37, Appl	C 952	15	1.5	3501	15	US-10-125-926A-37	Sequence 37, Appl
C 880	15	1.5	3501	14	US-10-146-726-37	Sequence 37, Appl	C 953	15	1.5	3501	15	US-10-125-930A-37	Sequence 37, Appl
C 881	15	1.5	3501	14	US-10-146-727-37	Sequence 37, Appl	C 954	15	1.5	3501	15	US-10-127-831A-37	Sequence 37, Appl
C 882	15	1.5	3501	14	US-10-146-788-37	Sequence 37, Appl	C 955	15	1.5	3501	15	US-10-127-837A-37	Sequence 37, Appl
C 883	15	1.5	3501	14	US-10-152-380-37	Sequence 37, Appl	C 956	15	1.5	3501	15	US-10-127-838B-37	Sequence 37, Appl
C 884	15	1.5	3501	14	US-10-153-934-37	Sequence 37, Appl	C 957	15	1.5	3501	15	US-10-127-842A-37	Sequence 37, Appl
C 885	15	1.5	3501	15	US-10-028-072-37	Sequence 37, Appl	C 958	15	1.5	3501	15	US-10-127-843A-37	Sequence 37, Appl
C 886	15	1.5	3501	15	US-10-121-049-37	Sequence 37, Appl	C 959	15	1.5	3501	15	US-10-127-845A-37	Sequence 37, Appl
C 887	15	1.5	3501	15	US-10-123-904-37	Sequence 37, Appl	C 960	15	1.5	3501	15	US-10-127-846A-37	Sequence 37, Appl
C 888	15	1.5	3501	15	US-10-140-470-37	Sequence 37, Appl	C 961	15	1.5	3501	15	US-10-127-848A-37	Sequence 37, Appl
C 889	15	1.5	3501	15	US-10-175-746-37	Sequence 37, Appl	C 962	15	1.5	3501	15	US-10-127-849A-37	Sequence 37, Appl
C 890	15	1.5	3501	15	US-10-176-918-37	Sequence 37, Appl	C 963	15	1.5	3501	15	US-10-127-850A-37	Sequence 37, Appl
C 891	15	1.5	3501	15	US-10-176-921-37	Sequence 37, Appl	C 964	15	1.5	3501	15	US-10-127-851A-37	Sequence 37, Appl
C 892	15	1.5	3501	15	US-10-137-865-37	Sequence 37, Appl	C 965	15	1.5	3501	15	US-10-128-684A-37	Sequence 37, Appl



RESULT 4  
US-10-274-266-1  
; Sequence 1, Application US/10274266  
; Publication No. US20030059893A1  
; GENERAL INFORMATION:  
; APPLICANT: LADUNGA, Steven et al.  
; TITLE OF INVENTION: Isolated Human Secreted Proteins, Nucleic Acid Molecules Encoding  
; TITLE OF INVENTION: Secreted Proteins, And Uses Thereof  
; FILE REFERENCE: CL000839DIV  
; CURRENT APPLICATION NUMBER: US/10/274,266  
; CURRENT FILING DATE: 2002-10-21  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1146  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-274-266-1

Query Match 2.0%; Score 19; DB 15; Length 1146;  
Best Local Similarity 100.0%; Pred. No. 7;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 471 TCTGGATACTGCTCGTTT 489  
Db 639 TCTGGATACTGCTCGTTT 657

RESULT 5  
US-09-731-872-15  
; Sequence 15, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US3.REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 15  
; LENGTH: 1755  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 101..517  
; NAME/KEY: sig\_peptide  
; LOCATION: 101..199  
; OTHER INFORMATION: Von Heijne matrix  
; OTHER INFORMATION: score 3.57613483592743  
; OTHER INFORMATION: seq FLCIGMALCLRQA/TR  
US-09-731-872-15

Query Match 2.0%; Score 19; DB 11; Length 1755;  
Best Local Similarity 100.0%; Pred. No. 7.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 471 TCTGGATACTGCTCGTTT 489  
Db 737 TCTGGATACTGCTCGTTT 755

RESULT 6  
US-09-731-872-14

; Sequence 14, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US3.REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 14  
; LENGTH: 1759  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 101..1243  
; NAME/KEY: sig\_peptide  
; LOCATION: 101..199  
; OTHER INFORMATION: Von Heijne matrix  
; OTHER INFORMATION: score 3.57142340200611  
; OTHER INFORMATION: seq FLCIGMALCPRQA/TR  
US-09-731-872-14

Query Match 2.0%; Score 19; DB 11; Length 1759;  
Best Local Similarity 100.0%; Pred. No. 7.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 471 TCTGGATACTGCTCGTTT 489  
Db 739 TCTGGATACTGCTCGTTT 757

RESULT 7  
US-10-128-714-87/c  
; Sequence 87, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wenqi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
; TITLE OF INVENTION: Methods of Use  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 87  
; LENGTH: 2352  
; TYPE: DNA  
; ORGANISM: Aspergillus fumigatus  
US-10-128-714-87



Query Match 2.0%; Score 19; DB 15; Length 2352;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 CAAAGCAGAGGGTGAAAGC 174  
Db 2018 CAAAGCAGAGGGTGAAAGC 2000

RESULT 8  
US-10-128-714-5087/c  
; Sequence 5087, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wenqi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
; TITLE OF INVENTION: Methods of Use  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5087  
; LENGTH: 2352  
; TYPE: DNA  
; ORGANISM: Aspergillus fumigatus  
US-10-128-714-5087

Query Match 2.0%; Score 19; DB 15; Length 2352;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 CAAAGCAGAGGGTGAAAGC 174  
Db 2018 CAAAGCAGAGGGTGAAAGC 2000

RESULT 9  
US-10-274-266-3  
; Sequence 3, Application US/10274266  
; Publication No. US20030059893A1  
; GENERAL INFORMATION:  
; APPLICANT: LADUNGA, Steven et al.  
; TITLE OF INVENTION: Isolated Human Secreted Proteins, Nucleic Acid Molecules Encoding  
; TITLE OF INVENTION: Secreted Proteins, And Uses Thereof  
; FILE REFERENCE: CL000839DIV  
; CURRENT APPLICATION NUMBER: US/10/274,266  
; CURRENT FILING DATE: 2002-10-21  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 3144  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-274-266-3

Query Match 2.0%; Score 19; DB 15; Length 3144;  
Best Local Similarity 100.0%; Pred. No. 7.6;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 471 TCTGGATACTGCTCGTTTT 489  
Db 1638 TCTGGATACTGCTCGTTTT 1656

RESULT 10  
US-09-918-995-11651  
; Sequence 11651, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 11651  
; LENGTH: 492  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-918-995-11651

Query Match 1.9%; Score 18; DB 12; Length 492;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 625 GAAAAAGCCCAACACATA 642  
Db 464 GAAAAAGCCCAACACATA 481

RESULT 11  
US-10-027-632-233560  
; Sequence 233560, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 233560  
; LENGTH: 559  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-233560

Query Match 1.9%; Score 18; DB 15; Length 559;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 810 TGATTATTTGGGTTGA 827  
|||||  
Db 491 TGATTATTTGGGTTGA 508

RESULT 12  
US-09-974-300-6602/c  
; Sequence 6602, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; APPLICANT: Clausen, Ib Groth  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; TITLE OF INVENTION: Expression  
; FILE REFERENCE: 10085.500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/279,526  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 8481  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 6602  
; LENGTH: 828  
; TYPE: DNA  
; ORGANISM: Bacillus clausii  
US-09-974-300-6602

Query Match 1.9%; Score 18; DB 11; Length 828;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 346 AATGTTGCCAAATATGCT 363  
|||||  
Db 291 AATGTTGCCAAATATGCT 274

RESULT 13  
US-09-938-842A-4346  
; Sequence 4346, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 4346  
; LENGTH: 996  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-4346

Query Match 1.9%; Score 18; DB 11; Length 996;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 912 AGAATTGATGGAATAATC 929  
|||||  
Db 258 AGAATTGATGGAATAATC 275

RESULT 14  
US-09-938-842A-5127/c  
; Sequence 5127, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 5127  
; LENGTH: 996  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-5127

Query Match 1.9%; Score 18; DB 11; Length 996;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 912 AGAATTGATGGAATAATC 929  
|||||  
Db 739 AGAATTGATGGAATAATC 722

RESULT 15  
US-10-027-632-259728  
; Sequence 259728, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 259728  
; LENGTH: 2338  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-259728

Query Match 1.9%; Score 18; DB 15; Length 2338;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 628 AAAGCCAACCAATATTT 645  
| | | | | | | | | | | | | | | | | |  
Db 1734 AAAGCCAACCAATATTT 1751

Search completed: July 31, 2003, 13:50:22  
Job time : 313 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 11:01:27 ; Search time 1720 Seconds  
(without alignments)  
9152.333 Million cell updates/sec

Title: US-09-992-430B-21  
Perfect score: 972  
Sequence: 1 atgttccaagatacaaaagtc.....ttgagcatctggagatcaat 972

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_vrt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_mam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	2.2	520	10 AW239816	AW239816 pt11c.pk0
2	20	2.1	202	12 BE859322	BE859322 UI-M-AQ0-
3	20	2.1	578	9 AA460603	AA460603 zx60G08.r
4	20	2.1	580	17 BH554885	BH554885 BOGDC32TF
5	20	2.1	583	9 AA453969	AA453969 zx45C06.r
6	20	2.1	592	17 PT005G13U	AL446361 Parameciu



80	19	2.0	699	12	BE908297	BE908297 601500589	153	18	1.9	495	9	AU242478	AU242478
81	19	2.0	725	12	BG028730	BG028730 602292916	C 154	18	1.9	496	9	AI958031	AI958031 fc89h05.y
82	19	2.0	741	17	BH470939	BH470939 BOGKZ01TR	155	18	1.9	500	9	AU088428	AU088428 AU088428
83	19	2.0	743	9	AJ450618	AJ450618 AJ450618	156	18	1.9	503	9	AL365864	AL365864 McBA02H01
84	19	2.0	743	14	BQ424840	BQ424840 AGENCOURT	C 157	18	1.9	503	17	BH352659	BH352659 CH230-81J
85	19	2.0	745	13	BG915424	BG915424 602815628	C 158	18	1.9	506	9	AI140336	AI140336 qa81e12.x
86	19	2.0	755	14	BQ435491	BQ435491 AGENCOURT	159	18	1.9	507	14	BM883476	BM883476 rc07c08.y
87	19	2.0	762	14	BM803372	BM803372 AGENCOURT	160	18	1.9	510	12	BG516069	BG516069 EtESTed48
88	19	2.0	762	17	AZ199368	AZ199368 SP 1039_B	161	18	1.9	511	13	BI745273	BI745273 xk98d05.y
89	19	2.0	771	9	AI420467	AI420467 tf07f11.x	C 162	18	1.9	513	10	AW150267	AW150267 xg49h11.x
90	19	2.0	793	13	BJ130459	BJ130459 BJ130459	C 163	18	1.9	515	14	BQ490600	BQ490600 77-E01178
91	19	2.0	800	10	BE391605	BE391605 601287014	164	18	1.9	516	14	BM883508	BM883508 rc07f05.y
92	19	2.0	807	9	AI121147	AI121147 ue89h07.y	C 165	18	1.9	518	10	AV965836	AV965836 AV965836
93	19	2.0	825	13	BI857657	BI857657 603384975	166	18	1.9	522	13	BM338184	BM338184 MEST222-H
94	19	2.0	829	12	BF034561	BF034561 601455476	167	18	1.9	538	17	TA358D10P	TA358D10P T. bruce1
95	19	2.0	836	9	AU125410	AU125410 AU125410	168	18	1.9	540	17	AQ504874	AQ504874 RPCI-11-3
96	19	2.0	852	17	AZ192066	AZ192066 SP 1020_B	C 169	18	1.9	552	17	AZ908153	AZ908153 RPCI-24-1
97	19	2.0	853	13	BI33253	BI33253 602996593	170	18	1.9	554	17	AQ500945	AQ500945 V29F6 mTn
98	19	2.0	858	17	BH720064	BH720064 BOHZJ32TF	C 171	18	1.9	560	13	BM190310	BM190310 POSM01000
99	19	2.0	859	12	BF577140	BF577140 602135427	172	18	1.9	562	9	AA037768	AA037768 zk38e01.s
100	19	2.0	864	12	BF969541	BF969541 602271873	173	18	1.9	581	13	BI814536	BI814536 PfESToa2
101	19	2.0	888	14	BQ217137	BQ217137 AGENCOURT	C 174	18	1.9	581	17	AZ880562	AZ880562 RPCI-23-1
102	19	2.0	899	14	BQ876592	BQ876592 AGENCOURT	175	18	1.9	591	17	BH103018	BH103018 RPCI-24-3
103	19	2.0	925	13	BI912675	BI912675 603176322	176	18	1.9	601	9	AA660498	AA660498 00384 MCR
104	19	2.0	937	14	BQ929045	BQ929045 AGENCOURT	C 177	18	1.9	607	13	BJ275391	BJ275391 BJ275391
105	19	2.0	950	12	BG289481	BG289481 602384329	178	18	1.9	608	14	BQ596074	BQ596074 PfESToa2
106	19	2.0	1003	14	BQ062301	BQ062301 AGENCOURT	C 179	18	1.9	612	17	AZ413764	AZ413764 IM0197H06
107	19	2.0	1010	17	CNS05FKI	AL335115 Tetraodon	180	18	1.9	615	12	BG587734	BG587734 EST489509
108	19	2.0	1047	13	BM478115	BM478115 AGENCOURT	181	18	1.9	620	14	W58369	W58369 zd24c12.r1
109	19	2.0	1069	13	BM557491	BM557491 AGENCOURT	182	18	1.9	625	17	TA1E10Q	TA1E10Q T. bruce1
110	19	2.0	1119	13	BM543065	BM543065 AGENCOURT	183	18	1.9	625	17	AQ653087	AQ653087 Sheared D
111	19	2.0	1138	17	B11122	B11122 F16L17-T7 I	184	18	1.9	629	17	AG153190	AG153190 Pan trogl
112	18	1.9	155	17	BH820991	BH820991 BACPP15-I	C 185	18	1.9	636	17	AZ383689	AZ383689 IM0141L23
113	18	1.9	161	17	BH883481	BH883481 BACPP15-I	C 186	18	1.9	640	10	BB610625	BB610625 BB610625
114	18	1.9	211	9	AV162076	BH83481 hw44h05.b	C 187	18	1.9	641	17	AQ324799	AQ324799 mgxb0019H
115	18	1.9	218	17	AQ249058	AV162076 AV162076	C 188	18	1.9	644	12	BF648209	BF648209 NF038G12E
116	18	1.9	227	10	BB074489	AQ249058 T1B12-Sp6	C 189	18	1.9	646	9	AU237048	AU237048 AU237048
117	18	1.9	228	17	BH832567	BB074489 BB074489	C 190	18	1.9	646	14	BQ553851	BQ553851 H4024F05-
118	18	1.9	260	17	CNS00VXC	BH832567 BACPP35-L	C 191	18	1.9	652	13	BJ348361	BJ348361 BJ348361
119	18	1.9	291	17	BH789421	AL092862 Arabidops	C 192	18	1.9	669	17	AG015527	AG015527 Homo sapi
120	18	1.9	308	10	AV300623	BH789421 SALK 0191	193	18	1.9	679	17	AG044939	AG044939 Pan trogl
121	18	1.9	312	14	Z33949	AV300623 AV300623	C 194	18	1.9	689	17	AQ956644	AQ956644 LERAL21TF
122	18	1.9	314	10	BB548876	Z33949 ATTS3007 Gr	C 195	18	1.9	690	17	AZ383805	AZ383805 IM0141P05
123	18	1.9	315	10	BB255599	BB548876 BB548876	C 196	18	1.9	701	14	BQ655135	BQ655135 NXRVO90 F
124	18	1.9	319	10	AW480082	BB255599 BB255599	C 197	18	1.9	707	17	AZ976860	AZ976860 2M0252H05
125	18	1.9	333	12	BF590415	AW480082 29515 MAR	C 198	18	1.9	718	17	AQ854086	AQ854086 nbxb0053B
126	18	1.9	335	17	BH752588	BF590415 naa37f06.	C 199	18	1.9	724	17	AQ329870	AQ329870 nbxb0045L
127	18	1.9	342	10	BB560598	BH752588 SALK 0188	C 200	18	1.9	731	10	AV384022	AV384022 AV384022
128	18	1.9	359	10	AW435544	BB560598 BB560598	C 201	18	1.9	742	17	AZ993687	AZ993687 2M0278N16
129	18	1.9	387	17	AQ205678	AW435544 74488 MAR	C 202	18	1.9	746	17	AZ973355	AZ973355 2M0247K03
130	18	1.9	393	12	BG561482	AQ205678 HS 3236 B	C 203	18	1.9	752	17	B90686	B90686 CIT-HSP-216
131	18	1.9	400	14	BQ293743	BG561482 EtESTed77	C 204	18	1.9	773	17	AQ746888	AQ746888 HS 2275 A
132	18	1.9	403	10	AW827576	BQ293743 1091021D0	C 205	18	1.9	777	17	BH770459	BH770459 LLMGtag22
133	18	1.9	433	10	AV855203	AW827576 hm22d07.x	C 206	18	1.9	783	14	BQ137631	BQ137631 NF038A11N
134	18	1.9	433	10	BE220251	AV855203 AV855203	C 207	18	1.9	786	9	AL566376	AL566376 AL566376
135	18	1.9	436	13	BI808842	BE220251 hv70a01.x	C 208	18	1.9	793	17	BH728594	BH728594 BOMEN24TF
136	18	1.9	436	14	BQ484483	BI808842 E004D08 O	C 209	18	1.9	803	17	BH604923	BH604923 BOGCS74TF
137	18	1.9	443	9	AI757170	BQ484483 pmp1c.pk0	C 210	18	1.9	816	9	AA979820	AA979820 MEST0-F3.
138	18	1.9	443	17	AQ149892	AI757170 EtESTea31	C 211	18	1.9	823	17	BH734004	BH734004 BOHYL80TR
139	18	1.9	447	17	B82423	AQ149892 HS 3192 B	C 212	18	1.9	832	12	BF168427	BF168427 601776018
140	18	1.9	454	9	AU163359	B82423 RPCI11-15I8	C 213	18	1.9	835	9	AJ452569	AJ452569 AJ452569
141	18	1.9	456	12	BG561895	AU163359 AU163359	C 214	18	1.9	837	17	AZ534802	AZ534802 ENTCS53TF
142	18	1.9	456	17	AQ894054	BG561895 EtESTee03	C 215	18	1.9	842	12	BG739872	BG739872 602630809
143	18	1.9	461	14	BQ244475	AQ894054 HS 2036 B	C 216	18	1.9	844	17	AZ681624	AZ681624 ENTLJ41TF
144	18	1.9	465	13	BI859457	BQ244475 TaE15036H	C 217	18	1.9	844	17	AZ687043	AZ687043 ENTHO85TF
145	18	1.9	466	9	AL693585	BI859457 603385301	C 218	18	1.9	850	17	BH541920	BH541920 BOHBW37TF
146	18	1.9	467	17	AQ624333	AL693585 AL693585	C 219	18	1.9	868	17	AZ667930	AZ667930 ENTGK16TF
147	18	1.9	473	12	BG725117	AQ624333 HS 2103 B	C 220	18	1.9	898	17	AZ530682	AL054974 Drosophil
148	18	1.9	474	12	BG561481	BG725117 EtESTed46	C 221	18	1.9	900	17	AZ531757	AZ531757 ENTBZ21TF
149	18	1.9	476	17	AQ462860	BG561481 EtESTed77	C 222	18	1.9	907	17	AZ670338	AZ670338 ENTIY89TF
150	18	1.9	482	12	BG517042	AQ462860 HS 5207 A	C 223	18	1.9	931	14	BQ895624	BQ895624 AGENCOURT
151	18	1.9	485	12	BF043063	BG517042 EtESTed69	C 224	18	1.9	968	12	BG208890	BG208890 RST28282
152	18	1.9	494	17	AQ078844	BF043063 BP250026B	C 225	18	1.9				
						AQ078844 CIT-HSP-2							

C 226	18	1.9	968	12	BG462141	BG462141 RST45155	299	17	1.7	367	17	BH881796	BH881796 hv28b03.g
C 227	18	1.9	972	17	CNS02EFO	AL193677 Tetraodon	C 300	17	1.7	368	9	AI470248	AI470248 tj17g05.x
C 228	18	1.9	986	14	BQ227634	BQ227634 AGENCOURT	C 301	17	1.7	369	13	BJ038491	BJ038491 BJ038491
C 229	18	1.9	989	12	BG869189	BG869189 602789072	C 302	17	1.7	370	17	AQ280064	AQ280064 CITBI-E1-
C 230	18	1.9	1009	13	BI830589	BI830589 603073120	C 303	17	1.7	372	9	AA818228	AA818228 UI-R-A0-a
C 231	18	1.9	1038	12	BF678823	BF678823 602153126	C 304	17	1.7	372	9	AV186746	AV186746 AV186746
C 232	18	1.9	1068	12	BF531186	BF531186 602091174	C 305	17	1.7	373	9	AA252795	AA252795 zb27a12.s
C 233	18	1.9	1114	12	BG547370	BG547370 602574778	C 306	17	1.7	373	17	AQ563048	AQ563048 HS_5301_A
C 234	18	1.9	1376	12	BE744646	BE744646 601577484	C 307	17	1.7	374	17	AZ430751	AZ430751 1M0215D20
C 235	18	1.9	1522	11	AK005882	AK005882 Mus muscu	C 308	17	1.7	375	9	AI029427	AI029427 UI-R-C0-i
C 236	17	1.7	138	17	AQ079932	AQ079932 CIT-HSP-2	C 309	17	1.7	375	14	C42506	C42506 C42506 Yuji
C 237	17	1.7	144	12	BE764667	BE764667 PM2-NT007	C 310	17	1.7	382	9	AA046522	AA046522 zk62f04.s
C 238	17	1.7	178	12	BF771253	BF771253 RC2-IT004	C 311	17	1.7	385	14	T97417	T97417 ye53h01.r1
C 239	17	1.7	178	12	BF920486	BF920486 QV0-NT010	C 312	17	1.7	386	10	AW357152	AW357152 40021 MAR
C 240	17	1.7	193	9	AA060952	AA060952 mj80h12.r	C 313	17	1.7	388	13	BI121099	BI121099 F028P60Y
C 241	17	1.7	193	10	BB213891	BB213891 BB213891	C 314	17	1.7	391	10	AW360330	AW360330 47894 MAR
C 242	17	1.7	193	14	H34245	H34245 EST110974 R	C 315	17	1.7	392	17	AQ003737	AQ003737 CpG0363A
C 243	17	1.7	196	12	BG629287	BG629287 CC-esf1cL	C 316	17	1.7	395	17	AZ715020	AZ715020 RPCI-24-1
C 244	17	1.7	198	13	BI643414	BI643414 RS1_B08 S	C 317	17	1.7	397	9	AA991355	AA991355 os51e12.s
C 245	17	1.7	224	10	BE051426	BE051426 za82f12.g	C 318	17	1.7	397	13	BJ070348	BJ070348 BJ070348
C 246	17	1.7	229	12	BG139606	BG139606 EST479964	C 319	17	1.7	398	17	AQ845497	AQ845497 LMAJFV1_1
C 247	17	1.7	235	10	BB065417	BB065417 BB065417	C 320	17	1.7	398	17	AQ127189	AQ127189 HS_3042_B
C 248	17	1.7	236	10	BB582827	BB582827 BB582827	C 321	17	1.7	400	17	BH881795	BH881795 hv28b03_b
C 249	17	1.7	237	10	BE142063	BE142063 CM3-HT013	C 322	17	1.7	402	12	BF892586	BF892586 QV1-MT013
C 250	17	1.7	254	12	BF920487	BF920487 QV0-NT010	C 323	17	1.7	402	17	B74335	B74335 CIT-HSP-204
C 251	17	1.7	256	10	AV440385	AV440385 AV440385	C 324	17	1.7	404	13	BI325478	BI325478 949039F10
C 252	17	1.7	263	17	AZ323839	AZ323839 1M0052F12	C 325	17	1.7	405	12	BE863227	BE863227 UI-M-BH0-
C 253	17	1.7	264	13	BI626279	BI626279 RH66849.5	C 326	17	1.7	406	17	AQ544703	AQ544703 CITBI-E1-
C 254	17	1.7	264	17	AQ350411	AQ350411 RPCI11-12	C 327	17	1.7	413	13	BI130057	BI130057 G101P77Y
C 255	17	1.7	277	12	BG139601	BG139601 EST479959	C 328	17	1.7	415	17	AF192545	AF192545 AF192545
C 256	17	1.7	278	12	BF920484	BF920484 QV0-NT010	C 329	17	1.7	416	10	AW379808	AW379808 RC3-HT025
C 257	17	1.7	282	10	BB046375	BB046375 BB046375	C 330	17	1.7	416	12	BF198160	BF198160 247009 MA
C 258	17	1.7	286	12	BG520462	BG520462 ps16h07.y	C 331	17	1.7	417	17	BH876508	BH876508 hr31e06.g
C 259	17	1.7	289	9	AV259207	AV259207 AV259207	C 332	17	1.7	418	9	AI299131	AI299131 qn14h10.x
C 260	17	1.7	290	14	BQ164880	BQ164880 EST610749	C 333	17	1.7	419	17	AQ338183	AQ338183 HS_2215_B
C 261	17	1.7	291	9	AV114792	AV114792 AV114792	C 334	17	1.7	420	10	BE192157	BE192157 db88b10.x
C 262	17	1.7	291	17	AQ116868	AQ116868 HS_2255_A	C 335	17	1.7	421	10	AW871992	AW871992 da95g02.y
C 263	17	1.7	292	10	BB359000	BB359000 BB359000	C 336	17	1.7	421	14	BQ637637	BQ637637 he12f09.y
C 264	17	1.7	307	14	BQ168387	BQ168387 WHE2102 D	C 337	17	1.7	425	12	BG629736	BG629736 CC-esf1cL
C 265	17	1.7	308	10	AW086543	AW086543 ga01g05.x	C 338	17	1.7	425	14	BQ129202	BQ129202 ij34f08.x
C 266	17	1.7	308	13	BM153847	BM153847 168 Giant	C 339	17	1.7	426	9	AI179145	AI179145 EST222829
C 267	17	1.7	311	10	AW098821	AW098821 ga01g05.y	C 340	17	1.7	427	17	AQ075054	AQ075054 CIT-HSP-2
C 268	17	1.7	312	9	AI031510	AI031510 16H8M Xen	C 341	17	1.7	427	17	AQ845498	AQ845498 LMAJFV1_1
C 269	17	1.7	313	9	AI511375	AI511375 UI-R-C3-s	C 342	17	1.7	427	17	BH613795	BH613795 SALK_0349
C 270	17	1.7	314	9	AA774997	AA774997 ae84f09.s	C 343	17	1.7	428	14	BQ629594	BQ629594 saq04a03.
C 271	17	1.7	316	9	AA631221	AA631221 nq80h02.s	C 344	17	1.7	431	10	AW641799	AW641799 cml1e04.w
C 272	17	1.7	318	9	AI249676	AI249676 qx50f11.x	C 345	17	1.7	434	9	AI165361	AI165361 A082p22u
C 273	17	1.7	322	13	BJ499046	BJ499046 BJ499046	C 346	17	1.7	434	14	BQ451614	BQ451614 PfESToab0
C 274	17	1.7	324	14	BQ455252	BQ455252 ke11b12.y	C 347	17	1.7	436	10	AW326143	AW326143 18298 MAR
C 275	17	1.7	328	10	BB377800	BB377800 BB377800	C 348	17	1.7	436	17	AQ005448	AQ005448 CIT-HSP-2
C 276	17	1.7	329	14	BQ455106	BQ455106 ke18e02.y	C 349	17	1.7	438	9	AA730822	AA730822 nw49e05.s
C 277	17	1.7	329	17	CNS00UDG	AL090850 Arabidops	C 350	17	1.7	438	9	AI167138	AI167138 oo08a07.x
C 278	17	1.7	332	14	Z45335	Z45335 HSC2MH071 n	C 351	17	1.7	442	9	AI783612	AI783612 c6d11a1.r
C 279	17	1.7	334	13	BJ059324	BJ059324 BJ059324	C 352	17	1.7	443	9	AI356577	AI356577 qy15g08.x
C 280	17	1.7	337	17	AQ591086	AQ591086 HS_5391_B	C 353	17	1.7	444	14	BQ129332	BQ129332 ij34f08.y
C 281	17	1.7	341	10	BE072577	BE072577 PM0-BT054	C 354	17	1.7	444	17	AQ456916	AQ456916 HS_5153_A
C 282	17	1.7	343	13	BM516501	BM516501 kj55b04.y	C 355	17	1.7	445	10	BE606644	BE606644 WHE0901_F
C 283	17	1.7	346	13	BM320163	BM320163 rs85h11.y	C 356	17	1.7	445	12	BF911135	BF911135 PM1-UT006
C 284	17	1.7	349	14	BQ283143	BQ283143 WHE33086 F	C 357	17	1.7	446	9	AI458373	AI458373 tj94e12.x
C 285	17	1.7	350	12	BF370475	BF370475 QV0-FN006	C 358	17	1.7	446	9	AA542638	AA542638 fa08e01.r
C 286	17	1.7	350	17	AZ315331	AZ315331 1M0032F05	C 359	17	1.7	447	9	AA212259	AA212259 mu78b04.r
C 287	17	1.7	351	13	BI816330	BI816330 PfESToaa3	C 360	17	1.7	448	10	AV874288	AV874288 AV874288
C 288	17	1.7	352	14	BQ282957	BQ282957 WHE33084_B	C 361	17	1.7	449	12	BF250424	BF250424 pa98h11.y
C 289	17	1.7	353	9	AI974661	AI974661 Tl13116e_	C 362	17	1.7	450	10	AW159469	AW159469 za99f05.x
C 290	17	1.7	353	17	BH267730	BH267730 CH230-215	C 363	17	1.7	451	13	BI743101	BI743101 kx38f10.y
C 291	17	1.7	356	13	BM403451	BM403451 z4m4369 Z	C 364	17	1.7	451	17	AZ123749	AZ123749 RPCI-23-4
C 292	17	1.7	358	9	AA639002	AA639002 ns01e07.s	C 365	17	1.7	451	17	AQ130356	AQ130356 HS_3014_A
C 293	17	1.7	358	12	BF868657	BF868657 IL3-ET011	C 366	17	1.7	452	10	AW088472	AW088472 xd27d03.x
C 294	17	1.7	358	17	BH58474	BH58474 B5_5900d_	C 367	17	1.7	452	17	B97017	B97017 F1SLITFB IG
C 295	17	1.7	360	9	AV191206	AV191206 AV191206	C 368	17	1.7	454	12	BG018385	BG018385 daa23f09.
C 296	17	1.7	360	9	AV193616	AV193616 AV193616	C 369	17	1.7	454	17	AQ457675	AQ457675 HS_5127_B
C 297	17	1.7	360	14	C39347	C39347 C39347 Yuji	C 370	17	1.7	456	9	AU037199	AU037199 AU037199
C 298	17	1.7	366	9	AI596369	AI596369 me61d11.x	C 371	17	1.7	456	14	BQ562226	BQ562226 H4074G09--

C 372	17	1.7	457	9	AA819830	AA819830 UI-R-A0-a	445	17	1.7	523	10	AW642062	AW642062 cml14e07.w
373	17	1.7	457	9	AI974580	AI974580 T113030e	446	17	1.7	523	13	BJ174044	BJ174044 BJ174044
374	17	1.7	458	17	AQ518425	AQ518425 HS_5101.A	C 447	17	1.7	523	14	BQ457070	BQ457070 ke34e03.y
375	17	1.7	460	10	BE326460	BE326460 hw04b08.x	C 448	17	1.7	525	13	BI591864	BI591864 RH07665.5
C 376	17	1.7	460	13	BI554530	BI554530 603235660	C 449	17	1.7	527	13	BM388382	BM388382 UI-R-CN1-
377	17	1.7	460	17	AZ235235	AZ235235 RPCI-23-8	C 450	17	1.7	528	10	AW852541	AW852541 PM1-CT024
C 378	17	1.7	461	9	AA036753	AA036753 zk31f04.x	451	17	1.7	528	12	BG363904	BG363904 dc72b08.y
379	17	1.7	462	10	AW665386	AW665386 hi90b03.x	452	17	1.7	528	12	BF325329	BF325329 su33c07.y
380	17	1.7	462	14	N68492	N68492 zal14h04.s1	C 453	17	1.7	528	17	TA240D09Q	TA240D09Q
381	17	1.7	466	9	AA680607	AA680607 SWOV13CAN	C 454	17	1.7	529	9	AI2333785	AI2333785 T. brucei
382	17	1.7	467	17	AQ331702	AQ331702 HS_5011.A	C 455	17	1.7	529	17	BH256826	BH256826 CH230-243
C 383	17	1.7	468	9	AI079328	AI079328 oz36h02.x	C 456	17	1.7	529	17	BH259406	BH259406 CH230-29E
384	17	1.7	468	9	AA261861	AA261861 zsl18h01.r	C 457	17	1.7	531	9	AI411934	AI411934 EST240228
385	17	1.7	469	12	BF549221	BF549221 UI-R-A0-a	458	17	1.7	531	12	BF215657	BF215657 601881136
C 386	17	1.7	469	17	B58350	B58350 CIT-HSP-201	459	17	1.7	529	17	BH256826	BH256826 CH230-243
387	17	1.7	470	13	BI168657	BI168657 RE09489.5	C 459	17	1.7	532	10	AV604646	AV604646 AV604646
C 388	17	1.7	470	17	AQ226728	AQ226728 HS_2023.B	460	17	1.7	532	17	AZ020835	AZ020835 RPCI-23-3
C 389	17	1.7	471	9	AB033121	AB033121 AB033121	461	17	1.7	533	9	AI693169	AI693169 wd68d02.x
390	17	1.7	471	10	BE251903	BE251903 601107550	462	17	1.7	533	10	AV552896	AV552896 AV552896
391	17	1.7	471	12	BG893270	BG893270 daa96b09.	C 463	17	1.7	533	10	AW262168	AW262168 xq62b12.x
392	17	1.7	476	9	AA485609	AA485609 zx90g06.x	464	17	1.7	533	17	AZ701136	AZ701136 RPCI-23-2
393	17	1.7	476	17	AQ492951	AQ492951 HS_5215.B	465	17	1.7	534	12	BG613244	BG613244 602641103
394	17	1.7	478	9	AI692611	AI692611 wd69h06.x	C 466	17	1.7	535	12	BG813587	BG813587 daf40f12.
C 395	17	1.7	478	14	BQ457134	BQ457134 ke35c07.y	C 467	17	1.7	537	14	BQ273095	BQ273095 sa015h11.
C 396	17	1.7	478	17	AZ522856	AZ522856 212pbB03	C 468	17	1.7	537	14	C85593	C85593 C85593 Mous
397	17	1.7	479	13	BJ196840	BJ196840 BJ196840	C 469	17	1.7	537	17	AZ852370	AZ852370 2M0155M20
398	17	1.7	480	10	AW907895	AW907895 ur94g02.y	470	17	1.7	537	17	AQ415076	AQ415076 RPCI-11-1
C 399	17	1.7	480	13	BM567674	BM567674 sak77c11.	471	17	1.7	539	12	BG023362	BG023362 dg40e07.x
400	17	1.7	481	17	BH325153	BH325153 CH230-45I	C 472	17	1.7	541	9	AI959628	AI959628 fdi2b03.x
C 401	17	1.7	482	12	BF939572	BF939572 7o47a06.x	C 473	17	1.7	541	9	AU160233	AU160233 AU160233
C 402	17	1.7	484	10	AV849884	AV849884 AV849884	474	17	1.7	541	13	BJ053741	BJ053741 BJ053741
C 403	17	1.7	484	17	AQ619709	AQ619709 HS_5185.A	475	17	1.7	541	17	AZ034253	AZ034253 RPCI-23-3
404	17	1.7	485	17	AQ368629	AQ368629 HS_5038.A	C 476	17	1.7	542	13	BM608588	BM608588 170006591
405	17	1.7	486	13	BM122098	BM122098 L0505E04-	477	17	1.7	543	13	BI578958	BI578958 RE72912.5
406	17	1.7	487	10	AW291737	AW291737 UI-H-B12-	C 478	17	1.7	545	14	BQ456234	BQ456234 ke28c01.y
C 407	17	1.7	488	12	BF600292	BF600292 264938.MA	479	17	1.7	545	17	AZ226488	AZ226488 RPCI-23-9
C 408	17	1.7	489	10	AW523387	AW523387 UI-R-B00-	C 480	17	1.7	545	17	BH095862	BH095862 RPCI-24-2
C 409	17	1.7	489	17	B47635	B47635 CIT-HSP-363	481	17	1.7	546	10	AV665672	AV665672 AV665672
C 410	17	1.7	490	10	BB854872	BB854872 BB854872	482	17	1.7	546	14	BM702640	BM702640 UI-E-CK1-
C 411	17	1.7	490	17	AQ992649	AQ992649 RPCI-23-3	483	17	1.7	549	10	AW921615	AW921615 EST352919
C 412	17	1.7	491	14	BQ027392	BQ027392 UI-H-C00-	484	17	1.7	550	13	BI846194	BI846194 fq78g04.y
413	17	1.7	492	12	BG234811	BG234811 dab64c10.	C 485	17	1.7	550	14	BQ456807	BQ456807 ke31c09.y
C 414	17	1.7	493	10	AV851209	AV851209 AV851209	C 486	17	1.7	550	17	BH876507	BH876507 hr31e06.b
C 415	17	1.7	493	12	BF010373	BF010373 SS98h10.y	C 487	17	1.7	550	17	AQ540692	AQ540692 RPCI-11-3
416	17	1.7	495	10	AV440069	AV440069 AV440069	488	17	1.7	552	17	AQ831889	AQ831889 HS_2066.B
417	17	1.7	498	10	BE148799	BE148799 CM4-HT024	C 489	17	1.7	553	13	BM269658	BM269658 sak02f03.
C 418	17	1.7	498	10	BE449104	BE449104 ut52e04.y	C 490	17	1.7	553	14	BM986667	BM986667 EST594261
419	17	1.7	504	13	BI634810	BI634810 RH39049.5	C 491	17	1.7	553	14	BQ455351	BQ455351 ke12f04.y
C 420	17	1.7	505	14	BQ451649	BQ451649 PFESTa0b0	C 492	17	1.7	554	13	BJ189003	BJ189003 BJ189003
C 421	17	1.7	505	17	AZ383966	AZ383966 1M0141117	C 493	17	1.7	554	14	BQ455913	BQ455913 ke24c01.y
422	17	1.7	506	12	BG050107	BG050107 FM1_66.G0	C 494	17	1.7	555	17	AZ112079	AZ112079 RPCI-23-4
C 423	17	1.7	507	17	AZ488795	AZ488795 1M0319F17	C 495	17	1.7	556	9	AI175578	AI175578 EST219133
C 424	17	1.7	507	17	AZ930600	AZ930600 474.dh256	C 496	17	1.7	556	14	BQ456174	BQ456174 ke27d02.y
425	17	1.7	508	10	AW732686	AW732686 bb11b12.y	C 497	17	1.7	557	13	BJ179654	BJ179654 BJ179654
426	17	1.7	509	14	R16858	R16858 yf85g01.r1	C 498	17	1.7	557	17	AQ745793	AQ745793 HS_2272.A
427	17	1.7	509	17	AQ319729	AQ319729 RPCI11-10	C 499	17	1.7	558	10	BE145224	BE145224 CM3-HT019
428	17	1.7	510	9	AA624724	AA624724 vn87d05.r	500	17	1.7	558	14	BQ285592	BQ285592 faa51g10.
C 429	17	1.7	513	10	AW755690	AW755690 sl06g11.y	501	17	1.7	560	9	AL588017	AL588017 AL588017
C 430	17	1.7	515	12	BG139419	BG139419 EST479861	C 502	17	1.7	562	14	BQ455450	BQ455450 ke14a04.y
C 431	17	1.7	515	12	BG139670	BG139670 EST480028	C 503	17	1.7	562	14	BQ456908	BQ456908 ke32e10.y
432	17	1.7	515	13	BI243071	BI243071 RE40619.5	C 504	17	1.7	562	14	BQ457137	BQ457137 ke35c10.y
C 433	17	1.7	515	13	BI301595	BI301595 UI-R-DL0-	505	17	1.7	563	12	BF554893	BF554893 UI-R-E0-c
434	17	1.7	516	10	BE292505	BE292505 601058010	C 506	17	1.7	563	14	BQ457407	BQ457407 ke38f08.y
435	17	1.7	517	9	AI543025	AI543025 SD09591.5	C 507	17	1.7	563	14	BQ457423	BQ457423 ke38h06.y
436	17	1.7	517	10	BE663192	BE663192 144201.MA	C 508	17	1.7	564	14	BQ563712	BQ563712 gi05h08.y
C 437	17	1.7	517	14	BQ457337	BQ457337 ke37h03.y	C 509	17	1.7	564	14	BQ455527	BQ455527 ke19g04.y
C 438	17	1.7	518	9	AA947873	AA947873 ok21h03.s	C 510	17	1.7	565	14	BQ455527	BQ455527 ke19g04.y
439	17	1.7	518	12	BF611895	BF611895 de89f04.y	C 511	17	1.7	565	14	BQ456025	BQ456025 ke25e05.y
C 440	17	1.7	518	17	CNS00V2J	AL091753 Arabidops	C 512	17	1.7	565	14	BQ456601	BQ456601 ke40g12.y
441	17	1.7	518	17	AQ402546	AQ402546 HS_5054.A	C 513	17	1.7	565	14	BQ456974	BQ456974 ke33d03.y
442	17	1.7	519	17	AQ425594	AQ425594 CITBI-E1-	C 514	17	1.7	565	14	BQ457082	BQ457082 ke34f04.y
C 443	17	1.7	522	9	AI232613	AI232613 EST229301	C 515	17	1.7	566	14	BQ455285	BQ455285 ke11f02.y
444	17	1.7	522	10	AW852543	AW852543 PM1-CT024	C 516	17	1.7	566	17	AQ251509	AQ251509 T6G4-SP6.
							C 517	17	1.7	567	10	AV679894	AV679894 AV679894



C 518	17	1.7	568	14	BQ454885	BQ454885	ke15g11.y	17	1.7	629	12	BF635500	BF635500	NF080A01D
C 519	17	1.7	568	14	BQ577311	BQ577311	pfESToab1	17	1.7	629	12	BE977052	BE977052	bs59c06.y
520	17	1.7	568	17	AZ093524	AZ093524	RPCI-23-4	17	1.7	631	9	AI109545	AI109545	GH08770.5
C 521	17	1.7	569	13	BM093998	BM093998	sah24c06.y	17	1.7	631	10	BB396007	BB396007	BB396007
C 522	17	1.7	569	14	BQ455549	BQ455549	ke20a04.y	17	1.7	631	17	AZ387055	AZ387055	1M0146018
523	17	1.7	569	14	BQ596264	BQ596264	pfESToab3	17	1.7	631	17	AQ201191	AQ201191	RPCI11-46
524	17	1.7	570	12	BG743895	BG743895	602722653	17	1.7	632	12	BG525671	BG525671	52-34 Ste
C 525	17	1.7	570	17	AQ608193	AQ608193	HS 2124.B	17	1.7	633	12	BG641299	BG641299	SD12618.5
C 526	17	1.7	571	17	BH088721	BH088721	RPCI-24-9	17	1.7	633	12	BG743874	BG743874	602722629
527	17	1.7	573	10	AW631512	AW631512	90686 MAR	17	1.7	634	14	BQ282406	BQ282406	WHE3071.F
C 528	17	1.7	574	12	BF288689	BF288689	EST453384	17	1.7	635	12	BG493801	BG493801	602542063
529	17	1.7	575	17	AZ812795	AZ812795	2M0079E17	17	1.7	636	14	BQ282418	BQ282418	WHE3071.G
C 530	17	1.7	576	10	AW381981	AW381981	QV4-HT031	17	1.7	637	12	BF126568	BF126568	601650823
C 531	17	1.7	576	14	BQ252737	BQ252737	san80b03.y	17	1.7	638	10	BB496928	BB496928	BB496928
532	17	1.7	577	13	BJ063698	BJ063698		17	1.7	638	12	BF399620	BF399620	UI-R-CA0-
533	17	1.7	577	17	AQ563856	AQ563856	HS 5338.B	17	1.7	638	17	AZ526996	AZ526996	262PbB07
C 534	17	1.7	579	9	AU019429	AU019429	AU019429	17	1.7	639	17	AZ232235	AZ232235	RPCI-23-5
C 535	17	1.7	579	10	AW174020	AW174020	fi38G07.y	17	1.7	642	9	AL789217	AL789217	AL789217
536	17	1.7	581	13	BI840514	BI840514	fs75B03.y	17	1.7	644	17	AZ339575	AZ339575	1M0071B12
C 537	17	1.7	581	17	BH350256	BH350256	CH230-198	17	1.7	645	17	B26457	B26457	F3G13TF IGF
C 538	17	1.7	582	17	B90812	B90812	CIT-HSP-216	17	1.7	646	13	BJ071062	BJ071062	BJ071062
C 539	17	1.7	582	17	AQ668753	AQ668753	HS 5435.A	17	1.7	647	12	BG522185	BG522185	19-26 Ste
C 540	17	1.7	583	14	BM986666	BM986666	EST594260	17	1.7	647	12	BG522319	BG522319	20-18 Ste
C 541	17	1.7	583	14	BQ454932	BQ454932	ke16d04.y	17	1.7	647	12	BG615596	BG615596	602642940
C 542	17	1.7	583	14	BQ455298	BQ455298	ke11h02.y	17	1.7	648	10	AV890204	AV890204	AV890204
C 543	17	1.7	583	14	BQ456550	BQ456550	ke40b08.y	17	1.7	648	14	BM860331	BM860331	fy37a02.y
C 544	17	1.7	583	17	AZ816503	AZ816503	2M0085H17	17	1.7	649	10	BB213848	BB213848	BB213848
C 545	17	1.7	583	17	BH873333	BH873333	hp44a10.g	17	1.7	650	9	AI533917	AI533917	SD05912.5
C 546	17	1.7	584	10	BE210610	BE210610	so50h10.y	17	1.7	650	10	BE555441	BE555441	sp88f12.y
C 547	17	1.7	584	17	AQ775182	AQ775182	HS 3151.A	17	1.7	650	12	BG526697	BG526697	62-95 Ste
548	17	1.7	585	17	AZ030920	AZ030920	RPCI-23-3	17	1.7	650	12	BF399637	BF399637	UI-R-CA0-
C 549	17	1.7	585	17	B22267	B22267	F18B7TR IGF	17	1.7	651	10	AV853734	AV853734	AV853734
C 550	17	1.7	587	10	AV859543	AV859543	AV859543	17	1.7	651	10	BB206241	BB206241	BB206241
C 551	17	1.7	587	10	AV881696	AV881696	AV881696	17	1.7	651	13	BM266319	BM266319	VL84 CDNA
C 552	17	1.7	587	12	BG526852	BG526852	40-59 Ste	17	1.7	652	10	AV882634	AV882634	AV882634
553	17	1.7	587	13	BI715210	BI715210	ic30c04.y	17	1.7	652	13	BI048391	BI048391	MR4-SN034
554	17	1.7	587	17	AZ093255	AZ093255	RPCI-23-4	17	1.7	655	17	B47767	B47767	RPCI11-2G9.
C 555	17	1.7	589	14	BQ455533	BQ455533	ke19g11.y	17	1.7	656	10	AV849072	AV849072	AV849072
C 556	17	1.7	589	14	BQ456678	BQ456678	ke41g08.y	17	1.7	657	10	AV852022	AV852022	AV852022
C 557	17	1.7	589	14	BQ577038	BQ577038	pfESToab1	17	1.7	657	17	AZ296672	AZ296672	RPCI-23-1
C 558	17	1.7	593	9	AA850598	AA850598	EST193366	17	1.7	658	13	BI635341	BI635341	SD16664.5
C 559	17	1.7	593	12	BG793540	BG793540	UTSW_SML3	17	1.7	658	17	BH652013	BH652013	BOMFZ37TF
C 560	17	1.7	594	9	AI103948	AI103948	EST213237	17	1.7	659	9	AL630304	AL630304	AL630304
C 561	17	1.7	594	13	BJ101439	BJ101439	BJ101439	17	1.7	660	13	BI167472	BI167472	RE07862.5
562	17	1.7	595	14	BM776425	BM776425	fy24F06.y	17	1.7	660	17	AZ571918	AZ571918	300PvB12
C 563	17	1.7	595	14	BQ456069	BQ456069	ke26a11.y	17	1.7	660	17	AG041165	AG041165	Pan trogl
564	17	1.7	597	17	BH784335	BH784335	fzmb013f0	17	1.7	661	17	AG133962	AG133962	Pan trogl
565	17	1.7	598	17	BH297447	BH297447	CH230-45E	17	1.7	663	9	AI386262	AI386262	mm48b10.y
C 566	17	1.7	599	10	AW562578	AW562578	660066F03	17	1.7	663	10	AV544569	AV544569	AV544569
C 567	17	1.7	599	17	AQ003687	AQ003687	CpG0332B	17	1.7	663	13	BI483694	BI483694	RE66617.5
C 568	17	1.7	599	17	CI22A1	CI22A1	Ciona int	17	1.7	663	14	BQ206933	BQ206933	UI-R-DZ1-
C 569	17	1.7	600	10	AW421413	AW421413	fj91a07.y	17	1.7	664	13	BI591883	BI591883	RH07690.5
570	17	1.7	602	12	BG525336	BG525336	48-80 Ste	17	1.7	668	17	AG081527	AG081527	Pan trogl
571	17	1.7	602	13	BI633124	BI633124	SD27115.5	17	1.7	669	12	BE823341	BE823341	GM700020B
572	17	1.7	604	12	BE823537	BE823537		17	1.7	670	17	AZ854753	AZ854753	2M0158E05
573	17	1.7	606	17	AQ393025	AQ393025	CITBI-E1-	17	1.7	671	12	BG094850	BG094850	uu81a02.x
574	17	1.7	607	10	AV861080	AV861080	AV861080	17	1.7	673	13	BJ279791	BJ279791	BJ279791
575	17	1.7	607	10	AV951274	AV951274	AV951274	17	1.7	675	17	AZ957643	AZ957643	2M0224006
576	17	1.7	607	12	BG710479	BG710479	pg11n.pk0	17	1.7	678	12	BE779030	BE779030	601464985
C 577	17	1.7	609	10	AW671599	AW671599	LG1_348.F	17	1.7	681	12	BG636062	BG636062	SD13651.5
C 578	17	1.7	612	17	AZ762041	AZ762041	1M0556N10	17	1.7	683	10	AW774027	AW774027	EST333013
579	17	1.7	613	12	BF239180	BF239180	601905580	17	1.7	683	12	BG143355	BG143355	1a96h12.x
580	17	1.7	616	13	BI826773	BI826773	603077258	17	1.7	683	17	AZ290835	AZ290835	RPCI-23-1
581	17	1.7	616	13	BM239182	BM239182	K0530C10-	17	1.7	684	14	BM943138	BM943138	UI-M-CG0p
C 582	17	1.7	616	17	AZ070892	AZ070892	RPCI-23-3	17	1.7	684	17	AG119117	AG119117	Pan trogl
583	17	1.7	618	9	AI134175	AI134175	GH11556.5	17	1.7	685	14	BQ245803	BQ245803	TaE15020A
C 584	17	1.7	618	17	AZ031420	AZ031420	RPCI-23-2	17	1.7	685	17	AG044115	AG044115	Pan trogl
585	17	1.7	621	12	BG709984	BG709984	pg11n.pk0	17	1.7	686	14	BQ407273	BQ407273	GA_EG010
C 586	17	1.7	623	17	DR8M23S	DR8M23S	Danio rer	17	1.7	687	12	BG723377	BG723377	602694036
C 587	17	1.7	625	12	BG525522	BG525522	50-77 Ste	17	1.7	688	10	AW565684	AW565684	LG1_348.F
C 588	17	1.7	625	17	AZ439966	AZ439966	1M0230N17	17	1.7	688	10	AW975680	AW975680	EST387789
589	17	1.7	626	17	AQ978056	AQ978056	RPCI-23-3	17	1.7	691	14	BQ407788	BQ407788	GA_EG000
C 590	17	1.7	627	17	BH760341	BH760341	BMBAC312D	17	1.7	693	9	AI517459	AI517459	GH28414.5



C 664	17	1.7	693	11	AK016337	Mus muscu	AK016337	17	1.7	792	14	BQ782146	UI-R-FFO-
C 665	17	1.7	695	12	BG744049	602722829	BG744049	17	1.7	796	10	BE131064	BE131064
C 666	17	1.7	695	13	BI204132	EST522172	BI204132	17	1.7	798	12	BG431308	BG431308
C 667	17	1.7	695	14	BQ575153	UI-H-EZ1-	BQ575153	17	1.7	798	13	BI825705	BI825705
C 668	17	1.7	695	14	BQ990952	QGF21122.	BQ990952	17	1.7	801	17	BH261954	BH261954
C 669	17	1.7	696	12	BF116218	7n79b04.x	BF116218	17	1.7	802	12	BF571948	BF571948
C 670	17	1.7	698	17	AG180368	Pan trogl	AG180368	17	1.7	809	9	AA768843	AA768843
C 671	17	1.7	698	17	AQ487070	RPCI-11-2	AQ487070	17	1.7	816	13	BI766386	BI766386
C 672	17	1.7	699	12	BG636315	SD13953.5	BG636315	17	1.7	817	10	BE270238	BE270238
C 673	17	1.7	699	17	AZ726406	RPCI-24-6	AZ726406	17	1.7	820	17	BH507597	BH507597
C 674	17	1.7	700	10	AV874972	AV874972	AV874972	17	1.7	822	14	BQ441248	BQ441248
C 675	17	1.7	701	13	BI240114	RE36763.5	BI240114	17	1.7	823	13	BM010546	BM010546
C 676	17	1.7	703	17	AZ838199	2M0133J18	AZ838199	17	1.7	826	17	AQ745631	AQ745631
C 677	17	1.7	704	13	BJ516044	BJ516044	BJ516044	17	1.7	828	12	BG744561	BG744561
C 678	17	1.7	705	13	BI226660	602951681	BI226660	17	1.7	828	13	BI311353	BI311353
C 679	17	1.7	706	10	BE264974	601194081	BE264974	17	1.7	829	17	BH539763	BH539763
C 680	17	1.7	707	17	BH293189	CH230-115	BH293189	17	1.7	831	12	BG744636	BG744636
C 681	17	1.7	707	17	AG125905	Pan trogl	AG125905	17	1.7	832	17	AZ186671	AZ186671
C 682	17	1.7	710	13	BI374883	RE62569.5	BI374883	17	1.7	833	13	BI916561	BI916561
C 683	17	1.7	710	13	BM592171	170006685	BM592171	17	1.7	839	9	AI068687	AI068687
C 684	17	1.7	710	17	B72080	CIT-HSP-363	B72080	17	1.7	840	13	BI462531	BI462531
C 685	17	1.7	712	13	BM657978	BM657978	BM657978	17	1.7	851	9	AU125099	AU125099
C 686	17	1.7	713	13	BI590975	BI590975	BI590975	17	1.7	851	17	AZ672130	AZ672130
C 687	17	1.7	713	17	AZ994333	2M0279F18	AZ994333	17	1.7	852	17	AQ749980	AQ749980
C 688	17	1.7	714	17	AZ956917	2M0223B12	AZ956917	17	1.7	856	13	BI107092	BI107092
C 689	17	1.7	717	13	BM658149	BM658149	BM658149	17	1.7	862	12	BF977573	BF977573
C 690	17	1.7	719	13	BM578500	BM578500	BM578500	17	1.7	864	9	AI789316	AI789316
C 691	17	1.7	721	12	BF032937	601455893	BF032937	17	1.7	866	12	BE785983	BE785983
C 692	17	1.7	722	10	AV850913	AV850913	AV850913	17	1.7	868	12	BF571981	BF571981
C 693	17	1.7	724	17	BH672553	BOHXZ35TR	BH672553	17	1.7	877	12	BG031457	BG031457
C 694	17	1.7	726	12	BG703971	602687076	BG703971	17	1.7	879	12	BF699164	BF699164
C 695	17	1.7	728	12	BG776861	602664044	BG776861	17	1.7	882	12	BG743090	BG743090
C 696	17	1.7	728	17	AQ960452	LERFD90TR	AQ960452	17	1.7	887	9	AA203338	AA203338
C 697	17	1.7	728	17	AG186153	Pan trogl	AG186153	17	1.7	890	12	BF247212	BF247212
C 698	17	1.7	729	10	BE561925	BE561925	BE561925	17	1.7	891	12	BF664359	BF664359
C 699	17	1.7	729	17	BH840362	LMCR20001	BH840362	17	1.7	891	17	BH159899	BH159899
C 700	17	1.7	730	17	AZ939224	2M0198B08	AZ939224	17	1.7	892	17	AZ541222	AZ541222
C 701	17	1.7	731	10	AV887784	AV887784	AV887784	17	1.7	893	17	AZ530663	AZ530663
C 702	17	1.7	734	12	BG546113	BG546113	BG546113	17	1.7	895	9	AL556437	AL556437
C 703	17	1.7	735	13	BI226898	BI226898	BI226898	17	1.7	895	17	AZ551104	AZ551104
C 704	17	1.7	736	9	AA698330	AA698330	AA698330	17	1.7	900	17	AZ530049	AZ530049
C 705	17	1.7	737	17	BH278203	BH278203	BH278203	17	1.7	905	12	BF106664	BF106664
C 706	17	1.7	738	13	BI930626	BI930626	BI930626	17	1.7	908	17	AZ530571	AZ530571
C 707	17	1.7	740	9	AI293274	GH16429.5	AI293274	17	1.7	914	9	AI106683	AI106683
C 708	17	1.7	740	9	AI587033	AI587033	AI587033	17	1.7	914	12	BE874186	BE874186
C 709	17	1.7	740	17	BH476015	BH476015	BH476015	17	1.7	915	9	AL578692	AL578692
C 710	17	1.7	744	14	W27445	W27445	W27445	17	1.7	916	17	BH149381	BH149381
C 711	17	1.7	746	13	BI461669	BI461669	BI461669	17	1.7	917	14	BQ717032	BQ717032
C 712	17	1.7	748	12	BE793507	BE793507	BE793507	17	1.7	919	14	BQ723759	BQ723759
C 713	17	1.7	750	17	AG079502	AG079502	AG079502	17	1.7	921	12	BG178537	BG178537
C 714	17	1.7	752	10	AV544760	AV544760	AV544760	17	1.7	923	9	AA203659	AA203659
C 715	17	1.7	754	12	BF607597	BF607597	BF607597	17	1.7	932	13	BI763601	BI763601
C 716	17	1.7	756	12	BG531305	BG531305	BG531305	17	1.7	933	17	BH147091	BH147091
C 717	17	1.7	763	12	BG621857	BG621857	BG621857	17	1.7	939	10	BE561026	BE561026
C 718	17	1.7	766	12	BG587712	BG587712	BG587712	17	1.7	941	12	BG031747	BG031747
C 719	17	1.7	767	17	AZ695195	AZ695195	AZ695195	17	1.7	941	17	AZ686088	AZ686088
C 720	17	1.7	767	17	BH040006	BH040006	BH040006	17	1.7	950	12	BG743957	BG743957
C 721	17	1.7	772	17	AZ989755	AZ989755	AZ989755	17	1.7	950	12	BE965277	BE965277
C 722	17	1.7	774	9	AJ454980	AJ454980	AJ454980	17	1.7	952	10	BE538509	BE538509
C 723	17	1.7	775	17	AQ749378	AQ749378	AQ749378	17	1.7	959	12	BF138512	BF138512
C 724	17	1.7	776	12	BG830544	BG830544	BG830544	17	1.7	968	17	CNS0435M	CNS0435M
C 725	17	1.7	777	9	AJ446862	AJ446862	AJ446862	17	1.7	974	14	BQ060869	BQ060869
C 726	17	1.7	777	13	BI643077	BI643077	BI643077	17	1.7	976	12	BG388484	BG388484
C 727	17	1.7	778	13	BI461648	BI461648	BI461648	17	1.7	993	12	BG285474	BG285474
C 728	17	1.7	778	17	BH419244	BH419244	BH419244	17	1.7	993	17	CNS02BQQ	CNS02BQQ
C 729	17	1.7	778	17	BH425379	BH425379	BH425379	17	1.7	999	12	BF662937	BF662937
C 730	17	1.7	779	9	AJ454745	AJ454745	AJ454745	17	1.7	1014	14	BQ056178	BQ056178
C 731	17	1.7	780	17	CNS00T88	CNS00T88	CNS00T88	17	1.7	1014	17	CNS03PPY	CNS03PPY
C 732	17	1.7	782	9	AI638303	AI638303	AI638303	17	1.7	1020	12	BG115731	BG115731
C 733	17	1.7	784	12	BF965634	BF965634	BF965634	17	1.7	1025	14	BQ050700	BQ050700
C 734	17	1.7	785	17	AQ872094	AQ872094	AQ872094	17	1.7	1032	12	BG752898	BG752898
C 735	17	1.7	786	14	BQ429272	BQ429272	BQ429272	17	1.7	1043	17	AG159941	AG159941
C 736	17	1.7	787	17	BH421909	BH421909	BH421909	17	1.7	1044	17	AZ691539	AZ691539

C 810	17	1.7	1054	17	CNS033WW	AL226697 Tetraodon	C 883	16	1.6	216	17	AZ627272	AZ627272 1M0469C03
C 811	17	1.7	1058	17	AZ209307 SP_0101.A	AZ209307 SP_0101.A	884	16	1.6	221	9	AA047149	AA047149 zk74f03.r
812	17	1.7	1072	14	BQ436660	BQ436660 AGENCOURT	885	16	1.6	221	10	AV316744	AV316744 AV316744
C 813	17	1.7	1079	17	CNS078FB	AL433917 T3 end of	886	16	1.6	221	12	BF327590	BF327590 MR0-BN011
C 814	17	1.7	1088	17	CNS0146I	AL103764 Drosophil	887	16	1.6	223	10	AV312977	AV312977 AV312977
815	17	1.7	1092	14	BM912276	BM912276 AGENCOURT	C 888	16	1.6	223	12	BF881863	BF881863 QV1-ET018
C 816	17	1.7	1101	17	CNS000F6	AL052843 Drosophil	C 889	16	1.6	224	17	BH887934	BH887934 LB01753a.
817	17	1.7	1101	17	CNS056TV	AL323788 Tetraodon	890	16	1.6	225	14	BU021305	BU021305 QHE29N15.
C 818	17	1.7	1101	17	CNS050Q6	AL346983 Tetraodon	C 891	16	1.6	225	17	AZ311542	AZ311542 1M0027I02
819	17	1.7	1148	13	BI561805	BI561805 603255241	892	16	1.6	226	14	BU022353	BU022353 QHE6K07.Y
C 820	17	1.7	1196	12	BE786921	BE786921 601477792	C 893	16	1.6	227	10	BB075920	BB075920 BB075920
C 821	17	1.7	1214	13	BM450200	BM450200 AGENCOURT	C 894	16	1.6	228	10	BB365780	BB365780 BB365780
822	17	1.7	1255	12	BG847688	BG847688 1024018F1	895	16	1.6	228	12	BG354388	BG354388 947036E02
823	17	1.7	1465	11	BC001557	BC001557 Homo sapi	896	16	1.6	228	14	BU018299	BU018299 QHE18D21.
824	17	1.7	1492	13	BI905819	BI905819 603062779	897	16	1.6	231	10	AV841514	AV841514 AV841514
C 825	17	1.7	1622	12	BG333619	BG333619 602460506	898	16	1.6	231	13	BM644704	BM644704 170006590
C 826	17	1.7	1677	11	BC008463	BC008463 Homo sapi	C 899	16	1.6	232	9	AI101061	AI101061 EST210350
C 827	17	1.7	1865	11	BC033014	BC033014 Homo sapi	C 900	16	1.6	232	10	AV378076	AV378076 AV378076
C 828	17	1.7	1913	12	BG029237	BG029237 602292451	C 901	16	1.6	232	10	AW968755	AW968755 EST380831
C 829	16	1.6	83	14	D12012	D12012 HUM000SI45	902	16	1.6	233	9	AV280040	AV280040 AV280040
C 830	16	1.6	83	14	D12013	D12013 HUM000SI46	903	16	1.6	233	12	BG315545	BG315545 PO3.0.292
C 831	16	1.6	97	9	AI453376	AI453376 tj37e01.x	904	16	1.6	234	9	AV283173	AV283173 AV283173
C 832	16	1.6	99	10	AV957678	AV957678 AV957678	C 905	16	1.6	235	14	BM896396	BM896396 952072F04
C 833	16	1.6	112	13	BM255274	BM255274 516964.MA	906	16	1.6	236	10	BE115330	BE115330 UI-R-BJ1-
834	16	1.6	121	14	BQ980017	BQ980017 QHI9014.Y	907	16	1.6	237	9	AV251581	AV251581 AV251581
C 835	16	1.6	123	10	AW603298	AW603298 RC1-CN001	908	16	1.6	237	10	AV352765	AV352765 AV352765
C 836	16	1.6	138	12	BF270504	BF270504 GA_EB000	C 909	16	1.6	237	10	AW404339	AW404339 UI-HF-BL0
C 837	16	1.6	139	12	BG063819	BG063819 H3011A10-	C 910	16	1.6	238	9	AV282534	AV282534 AV282534
C 838	16	1.6	142	9	AA786512	AA786512 m3a09a1.r	C 911	16	1.6	238	17	BH721259	BH721259 BOHSN39TF
C 839	16	1.6	147	10	AW813716	AW813716 RC3-ST019	C 912	16	1.6	239	9	AA054916	AA054916 SWAMCA150
C 840	16	1.6	147	13	BI245272	BI245272 949027A09	C 913	16	1.6	239	9	AA355859	AA355859 EST64322
C 841	16	1.6	152	9	AU254210	AU254210 AU254210	C 914	16	1.6	240	9	AA353180	AA353180 EST61278
C 842	16	1.6	152	12	BF073492	BF073492 220206.MA	C 915	16	1.6	240	12	BF843012	BF843012 RCS-HT103
C 843	16	1.6	152	17	AQ200890	AQ200890 RPCI11-62	916	16	1.6	240	14	BU020734	BU020734 QHE28D22.
C 844	16	1.6	162	17	BH857668	BH857668 SALK_0163	C 917	16	1.6	241	10	BB357717	BB357717 BB357717
845	16	1.6	163	17	AL770008	AL770008 Arabidops	C 918	16	1.6	241	10	BB484621	BB484621 BB484621
C 846	16	1.6	166	12	BE824540	BE824540 C0501A08-	919	16	1.6	241	14	BU019566	BU019566 QHE22H07.
847	16	1.6	170	17	B89920	B89920 CIT-HSP-217	C 920	16	1.6	243	10	AV340726	AV340726 AV340726
848	16	1.6	171	14	BQ974498	BQ974498 QHI15005.	921	16	1.6	243	14	BU016802	BU016802 QHE14A24.
849	16	1.6	172	10	AV380970	AV380970 AV380970	C 922	16	1.6	244	12	BG001036	BG001036 RCS-GN013
C 850	16	1.6	172	17	BH412959	BH412959 1007029H1	923	16	1.6	244	14	BQ871145	BQ871145 QGI11A02.
C 851	16	1.6	172	17	BH412960	BH412960 1007029H1	924	16	1.6	244	14	BU017924	BU017924 QHE17D19.
C 852	16	1.6	175	10	BB039886	BB039886 BB039886	925	16	1.6	244	17	AZ240676	AZ240676 RPCI-23-7
853	16	1.6	175	10	BE467206	BE467206 hz63a08.x	C 926	16	1.6	245	17	AZ442156	AZ442156 1M0234E15
C 854	16	1.6	175	13	BG981683	BG981683 MR3-CN014	C 927	16	1.6	246	10	AV838927	AV838927 AV838927
855	16	1.6	177	10	AW994704	AW994704 RC1-BN003	C 928	16	1.6	246	10	BB048783	BB048783 BB048783
C 856	16	1.6	178	10	AW786769	AW786769 120270.MA	929	16	1.6	246	13	BI200521	BI200521 n3e06f8.f
C 857	16	1.6	178	13	BG996646	BG996646 PM0-HT091	930	16	1.6	247	9	AA634877	AA634877 ab29c09.r
C 858	16	1.6	184	10	AW369462	AW369462 RC0-BN000	C 931	16	1.6	247	14	BU023018	BU023018 QHE9E24.Y
C 859	16	1.6	185	10	BB244210	BB244210 BB244210	C 932	16	1.6	248	9	AV284306	AV284306 AV284306
860	16	1.6	185	12	BG093690	BG093690 mab42c12.	933	16	1.6	248	14	F02957	F02957 HSCIHG032.n
C 861	16	1.6	189	10	BB533465	BB533465 BB533465	C 934	16	1.6	249	9	AV281976	AV281976 AV281976
C 862	16	1.6	192	14	BQ985255	BQ985255 QGEF08.Y	935	16	1.6	249	14	BQ911698	BQ911698 QHA17017.
C 863	16	1.6	192	14	N56521	N56521 LY1254F.Hum	C 936	16	1.6	250	9	AA378757	AA378757 EST91491
864	16	1.6	194	12	BG689903	BG689903 338226.BA	C 937	16	1.6	250	10	AW703139	AW703139 TGEStz283
865	16	1.6	195	10	BB037967	BB037967 BB037967	938	16	1.6	250	14	BQ864574	BQ864574 QGC27B09.
C 866	16	1.6	197	17	AZ113260	AZ113260 RPCI-23-9	939	16	1.6	251	9	AA090325	AA090325 Y0049.seq
C 867	16	1.6	199	10	AV347152	AV347152 AV347152	940	16	1.6	251	14	BM855216	BM855216 K-EST0138
C 868	16	1.6	200	10	AW325342	AW325342 16434.MAR	C 941	16	1.6	256	10	BB048063	BB048063 BB048063
869	16	1.6	201	9	AI926497	AI926497 wo45e01.x	C 942	16	1.6	257	10	AV354539	AV354539 AV354539
C 870	16	1.6	201	9	AA215843	AA215843 csh0134.s	C 943	16	1.6	257	10	BB011133	BB011133 BB011133
C 871	16	1.6	202	10	AW374613	AW374613 MRL1-CT005	C 944	16	1.6	257	10	BB150627	BB150627 BB150627
C 872	16	1.6	205	10	AW662398	AW662398 hh83h09.Y	945	16	1.6	257	10	BB473827	BB473827 BB473827
C 873	16	1.6	207	10	BB201027	BB201027 BB201027	946	16	1.6	257	14	F02250	F02250 HSC0TF082.n
874	16	1.6	208	14	BQ720994	BQ720994 AGENCOURT	C 947	16	1.6	258	14	BQ451014	BQ451014 PfESToab0
875	16	1.6	209	9	AI646472	AI646472 vv37g12.x	948	16	1.6	258	14	BU015991	BU015991 QHE11M16.
876	16	1.6	209	12	BF750353	BF750353 RC0-BN041	949	16	1.6	258	14	BU016786	BU016786 QHE14A05.
877	16	1.6	209	13	BM225219	BM225219 K0208C06-	950	16	1.6	258	14	BU018586	BU018586 QHE19B19.
878	16	1.6	209	14	BQ849199	BQ849199 QGA9D22.Y	951	16	1.6	259	9	AV285604	AV285604 AV285604
879	16	1.6	209	14	T30799	T30799 EST22863.Hu	952	16	1.6	259	14	BU018174	BU018174 QHE17007.
C 880	16	1.6	211	10	BE187111	BE187111 NXNV.159	C 953	16	1.6	260	10	BB349554	BB349554 BB349554
C 881	16	1.6	212	10	AV364654	AV364654 AV364654	954	16	1.6	260	10	BB584462	BB584462 BB584462
C 882	16	1.6	213	17	BH023760	BH023760 BG02096.D	955	16	1.6	260	14	BU018047	BU018047 QHE17123.

956	16	1.6	260	14	BU020476
957	16	1.6	260	14	BU021276
c 958	16	1.6	260	17	AZ407595
	16	1.6	261	13	BJ332879
959	16	1.6	261	14	BU017479
960	16	1.6	261	14	BU020404
961	16	1.6	261	14	BU020404
c 962	16	1.6	262	9	AV269236
	16	1.6	262	12	BE838615
963	16	1.6	262	14	BU023120
964	16	1.6	262	17	AL763404
c 965	16	1.6	262	10	BB438790
c 966	16	1.6	263	9	AV100873
c 967	16	1.6	264	14	BU022045
968	16	1.6	264	14	BU022045
969	16	1.6	265	9	AI488169
c 970	16	1.6	265	10	BE420050
971	16	1.6	265	14	BU021232
c 972	16	1.6	267	9	AV104868
	16	1.6	267	14	BQ849227
973	16	1.6	267	10	BB101858
974	16	1.6	268	10	BB157598
c 975	16	1.6	268	10	BB486516
c 976	16	1.6	268	14	BQ864085
977	16	1.6	268	14	BU029017
978	16	1.6	268	17	B36163
979	16	1.6	269	10	BB129126
c 980	16	1.6	269	14	BQ913583
981	16	1.6	269	14	BQ913773
982	16	1.6	270	14	BQ992516
983	16	1.6	271	9	AA236462
984	16	1.6	272	10	BB442069
c 985	16	1.6	272	10	BB552337
c 986	16	1.6	272	14	BQ845065
987	16	1.6	272	14	BQ862202
988	16	1.6	272	14	BQ865453
989	16	1.6	272	14	BQ865638
990	16	1.6	272	14	BQ865878
991	16	1.6	272	17	AZ923078
c 992	16	1.6	273	9	AV278627
c 993	16	1.6	273	10	BB456873
c 994	16	1.6	273	14	BQ980187
995	16	1.6	273	14	BQ986323
996	16	1.6	274	10	BB004442
c 997	16	1.6	274	10	BE644225
c 998	16	1.6	274	12	BG439113
c 999	16	1.6	274	14	BQ872420
1000	16	1.6	274	14	BQ872420

ALIGNMENTS

RESULT 1  
AW239816/c 520 bp mRNA linear EST 27-APR-2000  
LOCUS pt1lc.pk002.k24 chicken MDV infected T cell cDNA library Gallus  
DEFINITION gallus CDNA clone pt1lc.pk002.k24 5', mRNA sequence.

ACCESSION AW239816  
VERSION AW239816.1 GI:6579556  
KEYWORDS EST.  
SOURCE chicken.

ORGANISM

Gallus gallus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 520)

REFERENCE

AUTHORS Morgan, R.  
TITLE Chicken ESTs from MDV infected T cells  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robin Morgan  
University of Delaware  
Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-1341  
Fax: 302-831-2822  
Email: morgan@udel.edu

Clones can be ordered online at <http://www.chickest.udel.edu>.  
FEATURES  
source Location/Qualifiers  
1..520  
/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
/clone="pt1lc.pk002.k24"  
/clone\_lib="chicken MDV infected T cell cDNA library"  
/sex="Male and Female"  
/cell\_type="MDV infected splenic T cell"  
/lab\_host="E.coli DH5 alpha"  
/note="Vector: pB42AD"  
BASE COUNT 146 a 104 c 105 g 165 t  
ORIGIN

Query Match 2.2%; Score 21; DB 10; Length 520;  
Best Local Similarity 100.0%; Pred. No. 4.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 947 CAGTGATTGAGCATCTGGAGA 967  
|||||  
Db 192 CAGTGATTGAGCATCTGGAGA 172

RESULT 2

BE859322 202 bp mRNA linear EST 29-SEP-2000  
LOCUS UI-M-AQ0-aah-c-11-0-UI.r1 NIH BMAP\_MHI Mus musculus CDNA clone  
DEFINITION UI-M-AQ0-aah-c-11-0-UI 5', mRNA sequence.  
ACCESSION BE859322  
VERSION BE859322.1 GI:10375046  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 202)  
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mEST@mail.nih.gov

CDNA Library Preparation: M.B. Soares Lab Clone distribution:  
Researchers may obtain BMAP CDNA clones from RESEARCH GENETICS. It  
should be noted that Bento Soares is generating a small number of  
additional specialized non-redundant arrays of BMAP cDNAs whose  
availability will be considered under appropriate and limited  
collaborative arrangements  
Seq primer: M13 Reverse.

FEATURES

source

Location/Qualifiers

1..202  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-AQ0-aah-c-11-0-UI"  
/clone\_lib="NIH BMAP\_MHI"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: p77T3D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; The  
NIH BMAP\_MHI library is a non-normalized library  
constructed from mouse hippocampus. The tag is a string  
of 5 nucleotides present between the Not I site and the  
oligo-dT track. The library was constructed as described  
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806  
1996. Tissue provided by Ms. Annie Novakovich,  
Zivic-Miller Laboratories."

```

BASE COUNT      63 a      61 c      49 g      29 t
ORIGIN
Query Match      2.1%; Score 20; DB 12; Length 202;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      273 TAACCAAAAAAATGACAAA 292
      |||||||
Db      92 TAACCAAAAAAATGACAAA 111

RESULT 3
AA460603
LOCUS
DEFINITION
zx60d08.r1 Soares testis NHT Homo sapiens cDNA clone IMAGE:795855
5' similar to gb:U13680 L-LACTATE DEHYDROGENASE X CHAIN (HUMAN);,
mRNA sequence.
ACCESSION      AA460603
VERSION
KEYWORDS
SOURCE
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 323.
FEATURES
Location/Qualifiers
source
1..578
/organism="Homo sapiens"
/db_xref="GDB:6039437"
/db_xref="taxon:9606"
/clone="IMAGE:795855"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      165 a      99 c      142 g      172 t
ORIGIN
Query Match      2.1%; Score 20; DB 9; Length 578;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      547 ATTGGAGAACATGGTGATTC 566
      |||||||
Db      341 ATTGGAGAACATGGTGATTC 360
```

```

RESULT 4
BH554885
LOCUS
DEFINITION
BH554885 BOGDC32TF BOGD Brassica oleracea genomic clone BOGDC32, DNA
sequence.
ACCESSION      BH554885
VERSION
KEYWORDS
SOURCE
ORGANISM      Brassica oleracea.
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 580)
Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOGDC32TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
FEATURES
Location/Qualifiers
source
1..580
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGDC32"
/clone_lib="BOGD"
/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers."
BASE COUNT      191 a      107 c      161 g      121 t
ORIGIN
Query Match      2.1%; Score 20; DB 17; Length 580;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      898 GATGAGAAGGAAATAGAATT 917
      |||||||
Db      466 GATGAGAAGGAAATAGAATT 485

RESULT 5
AA453969
LOCUS
DEFINITION
AA453969 zx45c06.r1 Soares testis NHT Homo sapiens cDNA clone IMAGE:795178
5' similar to gb:U13680 L-LACTATE DEHYDROGENASE X CHAIN (HUMAN);,
mRNA sequence.
ACCESSION      AA453969
VERSION
KEYWORDS
SOURCE
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 583)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
```



Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 370.

FEATURES source  
1. .583  
/organism="Homo sapiens"  
/db\_xref="GDB:6038074"  
/db\_xref="taxon:9606"  
/clone="IMAGE:795178"  
/clone\_lib="Soares\_testis\_NHT"  
/sex="male"  
/lab\_host="DH10B"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from Clontech Laboratories  
, Inc., and primed with a Not I - oligo(dT) primer [5',  
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization to Cot5, and was  
constructed by Bento Soares and M. Fatima Bonaldo. "  
BASE COUNT 161 a 102 c 148 g 172 t  
ORIGIN

Query Match 2.1%; Score 20; DB 9; Length 583;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 547 ATTGGAGAACATGGTGATTC 566  
|||||  
Db 154 ATTGGAGAACATGGTGATTC 173

RESULT 6  
PT005G13U-  
LOCUS  
DEFINITION PT005G13U 592 bp DNA linear GSS 13-JUL-2001  
Paramaecium tetraurelia sequence M05D07u of the end of plasmid  
PT005G13, genomic survey sequence.  
ACCESSION AL446361  
VERSION AL446361.1 GI:11121969  
KEYWORDS  
SOURCE Paramaecium tetraurelia.  
ORGANISM Paramaecium tetraurelia.  
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;  
Paramaecium.  
REFERENCE 1 (bases 1 to 592)  
AUTHORS Keller,A.M. and Cohen,J.  
TITLE An indexed genomic library for Paramaecium complementation cloning  
JOURNAL J. Eukaryot. Microbiol. 47 (1), 1-6 (2000)  
MEDLINE 20114709  
PUBMED 10651287  
REFERENCE 2 (bases 1 to 592)  
AUTHORS Dessen,P., Zagulski,M., Gromadka,R., Plattner,H., Kissmehl,R.,  
Meyer,E., Betermier,M., Schultz,J.E., Linder,J., Pearlman,R.E.,  
Kung,C., Forney,J., Satir,B.H., Van Houten,J.L., Keller,A.M.,  
Froissard,M., Sperling,L. and Cohen,J.  
TITLE Paramaecium genome survey: a pilot project  
JOURNAL Trends Genet. 17 (6), 306-308 (2001)  
MEDLINE 21273563  
PUBMED 11377780  
REFERENCE 3 (bases 1 to 592)  
AUTHORS Gromadka,R. and Zagulski,M.  
TITLE Random sequencing of the Paramaecium macronuclear genome  
JOURNAL Unpublished  
REMARK Institute of Biochemistry and Biophysics, Polish Academy of  
Sciences, Warsaw, Poland  
REFERENCE 4 (bases 1 to 592)  
AUTHORS Cohen,J. and Sperling,L.  
TITLE Direct Submission  
JOURNAL Submitted (01-NOV-2000) Paramaecium Genome Survey Project, Centre de

Genetique Moleculaire, Centre National de la Recherche  
Scientifique, 91198 Gif-sur-Yvette, France. E-mail:  
sperling@cgm.cnrs-gif.fr  
The present survey of the Paramaecium tetraurelia macronuclear  
genome consists of end sequences of a library of random 4-12 kb  
fragments obtained by Sau3A partial digestion of macronuclear DNA  
cloned in the BamHI site of pBSIIKS-. See [4].  
Genes are predicted from matches to other sequences. For more  
information about this sequence or the Paramaecium Project, see  
http://paramaecium.cgm.cnrs-gif.fr.

FEATURES source  
1. .592  
/organism="Paramaecium tetraurelia"  
/macronuclear  
/strain="stock d4-2"  
/db\_xref="taxon:5888"  
BASE COUNT 179 a 92 c 80 g 241 t  
ORIGIN

Query Match 2.1%; Score 20; DB 17; Length 592;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 633 CAACCACATATTTGATCAGA 652  
|||||  
Db 173 CAACCACATATTTGATCAGA 192

RESULT 7  
BQ792707/c  
LOCUS  
DEFINITION BQ792707 607 bp mRNA linear EST 30-JUL-2002  
EST 8427 Red Grape berries Lambda Triplex2 Library Vitis vinifera  
cDNA clone CM005B02 3', mRNA sequence.  
ACCESSION BQ792707  
VERSION BQ792707.1 GI:22007673  
KEYWORDS  
SOURCE Vitis vinifera.  
ORGANISM Vitis vinifera  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Vitaceae; Vitis.  
REFERENCE 1 (bases 1 to 607)  
AUTHORS Abbal,P., Agasse,A., Ageorges,A., Atanassova,R., Barrieu,F.,  
Couture,C., Dedaldechamp,F., Delrot,S., Glissant,D., Grimplet,J.,  
Hamdi,S., Romieu,C. and Terrier,N.  
TITLE Generation of Expressed Sequence Tag from Grape Berry (skin, pulp  
or seeds) at Various Developmental Stages  
JOURNAL Unpublished (2002)  
COMMENT Contact: Hamdi S.  
UMR 619 - Equipe Biologie de la Vigne  
Universite de Bordeaux I, Institut National de la Recherche  
Agronomique  
71, Avenue Edouard Bourleaux, BP 81, 33883 Villenave D'Ornon Cedex,  
France  
Tel: 00-33-(0)5-57-12-25-50  
Fax: 00-33-(0)5-57-12-25-48  
Email: s.hamdi@bordeaux.inra.fr  
Seq primer: T7.

FEATURES source  
1. .607  
/organism="Vitis vinifera"  
/cultivar="Cabernet Sauvignon"  
/db\_xref="taxon:29760"  
/clone="CM005B02"  
/clone\_lib="Red Grape berries Lambda Triplex2 Library"  
/dev\_stage="Harvest - 15,5 weeks post-flowering"  
/note="Organ: Fruit without seeds; Vector: Lambda Triplex2  
; Site 1: SfiIA; Site 2: SfiIB; Oriented library"  
BASE COUNT 169 a 148 c 128 g 162 t  
ORIGIN

Query Match 2.1%; Score 20; DB 14; Length 607;  
Best Local Similarity 100.0%; Pred. No. 18;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 AAAAAAATGGACAAACAGG 297  
|||||  
Db 176 AAAAAAATGGACAAACAGG 157

RESULT 8  
AI638822/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AI638822 616 bp mRNA linear EST 27-APR-1999  
tt32g02.x1 NCI CGAP GC6 Homo sapiens CDNA clone IMAGE:2242514 3'  
similar to gb:U13680 L-LACTATE DEHYDROGENASE X CHAIN (HUMAN);, mRNA  
sequence.  
AI638822  
AI638822.1 GI:4691056  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 616)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -40UP from Gibco  
High quality sequence stop: 430.  
Location/Qualifiers  
1. .616  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2242514"  
/clone\_lib="NCI\_CGAP\_GC6"  
/tissue\_type="pooled germ cell tumors"  
/lab\_host="DH10B"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA  
from the normalized library NCI\_CGAP GC4 was prepared, and  
ss circles were made in vitro. Following HAP purification,  
this DNA was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from a pool  
of 5,000 clones made from the same library (cloneIDs  
1257096-1258631, 1469064-1470983, and 1475592-1476743).  
Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 180 a 143 c 91 g 201 t 1 others  
ORIGIN

Query Match 2.1%; Score 20; DB 9; Length 616;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 547 ATTGGAGAACATGGTGATTC 566  
|||||  
Db 570 ATTGGAGAACATGGTGATTC 551

RESULT 9  
AG148738/c  
LOCUS  
DEFINITION  
ACCESSION

AG148738 680 bp DNA linear GSS 09-JAN-2002  
Pan troglodytes DNA, clone: RP43-011A13.TJ, genomic survey  
sequence.  
AG148738

VERSION AG148738.1 GI:16678416  
KEYWORDS GSS.  
SOURCE Pan troglodytes male lymphocytes DNA, clone\_lib:RPCI-43 Chimpanzee  
Male BAC Library clone:RP43-011A13.TJ.  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
REFERENCE  
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,  
Totoki,Y., Watanabe,H. and Sakaki,Y.  
TITLE BAC end sequences of Library RPCI-43  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 680)  
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,  
Totoki,Y., Watanabe,H. and Sakaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail:chimbbs@gsc.riken.go.jp, URL:<http://hgp.gsc.riken.go.jp/>,  
Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the chimpanzee BAC library RPCI-43 This BAC  
end was generated during the R&D process and may have higher chance  
of clone tracking errors.  
PRIMERS  
Sequencing: TJ  
LIBRARY  
Vector : pBACe3.6  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI.  
Location/Qualifiers  
1. .680  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="RP43-011A13.TJ"  
/sex="male"  
/cell\_type="lymphocytes"  
/clone\_lib="RPCI-43 Chimpanzee Male BAC Library"  
BASE COUNT 213 a 124 c 119 g 223 t 1 others  
ORIGIN

Query Match 2.1%; Score 20; DB 17; Length 680;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 TTACCAAAAAAATGGACAA 291  
|||||  
Db 457 TTACCAAAAAAATGGACAA 438

RESULT 10  
BI827602  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BI827602 713 bp mRNA linear EST 04-OCT-2001  
603073568F1 NIH\_MGC\_119 Homo sapiens CDNA clone IMAGE:5165608 5',  
mRNA sequence.  
BI827602  
BI827602.1 GI:15939139  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 713)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)  
cDNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be



```
Db      547 ATTGGAGAACATGGTGATTC 566

RESULT 13
BQ797938/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BQ797938      745 bp      mRNA      linear      EST 30-JUL-2002
EST 6876 Ripening Grape berries Lambda Zap II Library Vitis
vinifera cDNA clone RT093E05 3', mRNA sequence.
BQ797938
BQ797938.1   GI:22012904
EST.
Vitis vinifera.
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Vitaceae; Vitis.
1 (bases 1 to 745)
Abbal,P., Agasse,A., Ageorges,A., Atanassova,R., Barrieu,F.,
Couture,C., Dedaldechamp,F., Delrot,S., Glissant,D., Grimplet,J.,
Hamdi,S., Romieu,C. and Terrier,N.
Generation of Expressed Sequence Tag from Grape Berry (skin, pulp
or seeds) at Various Developmental Stages
Unpublished (2002)
Contact: Romieu C.
Unite de Recherche des Produits de la Vigne
Institut National de la Recherche Agronomique
2, place Viala, 34 060 Montpellier Cedex 01, France
Tel: 00-33-(0)4-99-61-28-62
Fax: 00-33-(0)4-99-61-28-57
Email: romieu@ensam.inra.fr
Seq primer: T7.

FEATURES
            Location/Qualifiers
            1..745
            /organism="Vitis vinifera"
            /cultivar="Shiraz"
            /db_xref="taxon:29760"
            /clone="RT093E05"
            /dev_stage="Ripening stage"
            /note="Organ: Fruit; Vector: Lambda Zap II; Site 1: Eco RI
; Site 2: XhoI; Oriented library, construction described
in Generation of ESTs from grape Berry (skin, pulp or
seeds) at various developmental stages by Terrier,N.,
Ageorges,A., Abbal,P., Romieu,C. in J. Plant Physiol. 158
(12): 1575-83 2001"

BASE COUNT      186 a      192 c      169 g      198 t
ORIGIN

Query Match      2.1%; Score 20; DB 14; Length 745;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      278 AAAAAAATGGACAACAAGG 297
      ||||||||||||||||||
Db      114 AAAAAAATGGACAACAAGG 95

RESULT 14
BG723296
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BG723296      756 bp      mRNA      linear      EST 08-MAY-2001
602690928F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4823094 5',
mRNA sequence.
BG723296
BG723296.1   GI:14002483
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 756)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

REFERENCE
AUTHORS
TITLE
JOURNAL

Db      547 ATTGGAGAACATGGTGATTC 566

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10732 row: h column: 07
High quality sequence stop: 755.

FEATURES
            Location/Qualifiers
            1..756
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4823094"
            /clone_lib="NIH_MGC_97"
            /lab_host="DH10B"
            /note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: Sali-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT      194 a      136 c      193 g      233 t
ORIGIN

Query Match      2.1%; Score 20; DB 12; Length 756;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      547 ATTGGAGAACATGGTGATTC 566
      ||||||||||||||||||
Db      653 ATTGGAGAACATGGTGATTC 672

RESULT 15
BI828443
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BI828443      774 bp      mRNA      linear      EST 04-OCT-2001
603078209F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169698 5',
mRNA sequence.
BI828443
BI828443.1   GI:15939993
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 774)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11422 row: b column: 03
High quality sequence stop: 772.

FEATURES
            Location/Qualifiers
            1..774
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5169698"
            /clone_lib="NIH_MGC_119"
```



/tissue type="medulla"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: EcoRV (destroyed); RNA source normal medulla from  
anonymous male age 27. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.3 kb, insert size range  
0.9-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 013. Note:  
this is a NIH MGC Library."

BASE COUNT	206 a	142 c	189 g	237 t
ORIGIN				

```
Query Match      2.1%; Score 20; DB 13; Length 774;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 547 ATTGGAGAACATGGTGATTC 566  
|||||  
Db 648 ATTGGAGAACATGGTGATTC 667

Search completed: July 31, 2003, 11:58:57  
Job time : 1863 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 06:49:36 ; Search time 2765 Seconds  
(without alignments)  
10230.715 Million cell updates/sec

Title: US-09-992-430B-21  
Perfect score: 972  
Sequence: 1 atgttccaagatacaaaagtc.....ttgagcatctggagatcaat 972

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	265.4	27.3	30176	8	SPAC186	AL157811 S.pombe c
2	216.4	22.3	1423	1	BACLDHAA	M22305 B.megateriu
C 3	205.2	21.1	10029	1	AE006274	AE006274 Lactococc
4	201.2	20.7	11158	1	AE007540	AE007540 Clostridi
5	199.4	20.5	948	1	AY098994	AY098994 Clostridi
C 6	185.8	19.1	301450	1	AP003185	AP003185 Clostridi
7	183.8	18.9	969	6	AX142689	AX142689 Sequence
8	183.8	18.9	969	6	AX143661	AX143661 Sequence
C 9	183.8	18.9	3019	1	AF270315	AF270315 Staphyloc
C 10	183.8	18.9	3019	6	AX145633	AX145633 Sequence
11	183.8	18.9	4093	1	AF269753	AF269753 Staphyloc
12	183.8	18.9	4093	6	AX145071	AX145071 Sequence
13	183.6	18.9	1255	6	AX406983	AX406983 Sequence
14	176.6	18.2	1130	1	TMLDH	X74302 T.maritima
15	176.6	18.2	12762	1	AE001823	AE001823 Thermotog
16	173.2	17.8	6984	1	AF054624	AF054624 Lactobaci
17	171.6	17.7	1260	1	LSU26688	U26688 Lactobacill
C 18	171	17.6	301250	1	AP004830	AP004830 Staphyloc
C 19	169.4	17.4	297850	1	AP003137	AP003137 Staphyloc
C 20	169.4	17.4	341350	1	AP003365	AP003365 Staphyloc
21	166.8	17.2	290250	1	AP004822	AP004822 Staphyloc
22	166.8	17.2	298050	1	AP003129	AP003129 Staphyloc
23	166.8	17.2	349999	1	AP003358	AP003358 Staphyloc
24	164.6	16.9	1168	1	LBALDH	M76708 L.casei lac
25	164.6	16.9	1364	1	LBALLCT	D12591 L.casei gen
26	161.2	16.6	1199	1	BPSCTB	X55119 B. psychros
C 27	157.2	16.2	12584	1	AE006557	AE006557 Streptoco
28	156.6	16.1	1260	1	STRLCT	L42474 Streptococc
29	156.6	16.1	1292	1	STRLDH	M72545 Streptococc
30	155.8	16.0	945	6	E28918	E28918 Lactate deh
31	155.8	16.0	1065	6	AX065003	AX065003 Sequence
C 32	155.6	16.0	50734	1	AE014152	AE014152 Streptoco
33	155.4	16.0	978	1	LUU78637	U78637 Lactococcus
34	155.2	16.0	1086	1	BACLDHA	M19395 B.caldotena
35	155.2	16.0	1244	1	SPR6LDH	AJ000336 Streptoco
36	155.2	16.0	11614	1	AE008483	AE008483 Streptoco
37	155.2	16.0	77743	2	SPNEU1910	AL449932 Streptoco
38	154.2	15.9	942	6	AX123294	AX123294 Sequence
C 39	154.2	15.9	224650	1	AL596164	AL596164 Listeria
C 40	154.2	15.9	309400	6	AX127153	AX127153 Sequence
C 41	154.2	15.9	325651	1	AP005283	AP005283 Corynebac
C 42	154.2	15.9	349980	6	AX415067	AX415067 Sequence
C 43	154.2	15.9	349980	6	AX453571	AX453571 Sequence
44	153.6	15.8	1368	1	BACLCTLDH	M28336 B.caldotena
C 45	153.6	15.8	5278	6	BD003907	BD003907 Polynucle

ALIGNMENTS

RESULT 1  
SPAC186/c  
LOCUS S.pombe chromosome I cosmid c186.  
DEFINITION S.pombe chromosome I cosmid c186.  
ACCESSION AL157811  
VERSION AL157811.1 GI:7024417  
KEYWORDS 2-hydroxyacid dehydrogenase; D-isomer specific 2-hydroxyacid dehydrogenases domain; gene free region; l-asparaginase precursor; l-lactate dehydrogenase; mitochondrial transit peptide; pyruvate decarboxylase; subtelomeric region; Tf1-type LTR; Tf2-type LTR; Uncharacterized protein family UPF0016 domain.  
SOURCE Schizosaccharomyces pombe  
ORGANISM Schizosaccharomyces pombe  
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
Schizosaccharomycetales; Schizosaccharomycetaceae;

Schizosaccharomyces.  
1 (bases 1 to 30176)  
Wood,V., Rajandream,M.A., Barrell,B.G., Seeger,K. and Harris,D.  
Direct Submission  
Submitted (22-FEB-2000) European Schizosaccharomyces genome  
sequencing project, Sanger Centre, The Wellcome Trust Genome  
Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk  
Notes:  
Details of yeast sequencing at the Sanger Centre are available on  
the World Wide Web.  
(URL, [http://www.sanger.ac.uk/Projects/S\\_pombe/](http://www.sanger.ac.uk/Projects/S_pombe/))  
During 1995 to 1996 about 66% of S. pombe chromosome 1 was  
sequenced by the Sanger Centre. The sequencing of the S. pombe  
genome is now being continued with funding from The European  
Commission. Fourteen European sequencing laboratories, including  
the Sanger Centre, are participating in the project.  
Protein coding regions (CDS) have been predicted with the help of  
computer analysis using the Genefinder program in PomBase (an ACEDB  
database) with additional predictions for the branch-acceptor sites  
supplied by the program Sp3splice. CAUTION: It is possible that for  
any individual CDS we may have underestimated or overestimated the  
number of introns/exons or we may not have chosen the correct  
splice donor/acceptor sites.  
CDS are numbered using the following system eg SPBC25H2.01c. SP (S.  
pombe), B (chromosome 2), c25H2 (cosmid name), .01 (first CDS), c  
(complementary strand).  
The more significant matches with motifs in the PROSITE database  
are also included but some of these may be fortuitous.  
The length in codons is given for each CDS.  
IMPORTANT: This sequence MAY NOT be the entire insert of the  
sequenced clone. It may be shorter because we only sequence  
overlapping sections once, or longer, because we arrange for a  
small overlap between neighbouring submissions.  
Cosmid c186 is overlapped at the 5' end by cosmid c869, EMBL entry  
SPAC869, accession number AL132779.  
Location/Qualifiers  
1. .30176  
/organism="Schizosaccharomyces pombe"  
/strain="972h-"  
/db\_xref="taxon:4896"  
/chromosome="I"  
/map="IR"  
/clone="cosmid c186"  
/note="gene free region"  
1. .2929  
/note="nominal overlap with SPAC869 S. pombe chromosome 1"  
2929. .3909  
/gene="SPAC186.01"  
2929. .3909  
/gene="SPAC186.01"  
/note="SPAC186.01"  
/note="SPAC186.01, len:326, SIMILARITY:Schizosaccharomyces  
pombe., CAB69629, hypothetical serine/threonin repeat  
protein., (416 aa), fasta scores: opt: 334, E():2.6e-13,  
(33.3% identity in 261 aa)"  
/codon\_start=1  
/label=SPAC186.01  
/product="hypothetical serine/threonine rich protein"  
/protein\_id="CAB75865.1"  
/db\_xref="GI:7024418"  
/translation="MNVVKYIIFSFALAPLLLVNANTYNFTQIKRSVNOAVLESSQD  
TNSVGGEASTACPLFTTIYNTNGITPGTTIYPTISITSGVSSNNIDETSVSSEIIT  
STITTTITSGSGLYTTTITGQNTPDVTEVVIPITAGTFTTTLTSGSSYPVATTVTRA  
SGTQSGEVEVITPSCGSPENSHLKDNDKISPSYVMDPNAPVRTNGAGREGNMFA  
STNGDNEGLNLFYDSTIQRLVTCDCQRPSTVYIEDPIIGNGFSSAWNLIKNSDGIF  
TPVESRNNRPLHFHVDNNGRVWMTSQEYDTEVSSTDERNFRANDVTQLQY"  
complement(5199. .6197)  
/gene="SPAC186.02c"  
complement(5199. .6197)  
/gene="SPAC186.02c"  
/note="SPAC186.02c, len:332, SIMILARITY:Zymomonas  
mobilis., DDH\_ZYMO, 2-hydroxyacid dehydrogenase homolog,  
(331 aa), fasta scores: opt: 743, E():0, (47.6% identity

in 319 aa)"  
/codon\_start=1  
/label=SPAC186.02c  
/product="2-hydroxyacid dehydrogenase homolog"  
/protein\_id="CAB75866.1"  
/db\_xref="GI:7024419"  
/translation="MRVVLFSQSQSYDRGPFEEANKTFNHEIYYHNFSLNKDTVSLAGK  
AQWCVFVNDQVDADTLKALAENGVKLVALRCGGYNNVNLKAASEYKITVVHVPSYSP  
FAVSEFTVGLLLSLNRKIHRAIVRVREDDFNIVGLGCDIHGKTGVIGTGKIGSNVA  
KCFKMGFGCDVLAYDINPDKKLENYGVQFVEQNEVLKKADFLCLHCPLTPSTTHVNS  
DSLALMKKGVTIVNTRGGIDITKALVDIDAIDSGQVGGCAIDVEGERNLFYKDLSEV  
IKDSTFQRLVNFNPVLVTSQAFFTEALCSIAHTLKSASDFYTNLSDESVIANK"  
complement(5310. .5966)  
/gene="SPAC186.02c"  
/note="Match to PF00389 2-Hacid DH, D-isomer specific  
2-hydroxyacid dehydrogenases Score 325.65"  
6198. .10434  
/note="gene free region"  
10435. .11517  
/gene="SPAC186.03"  
10435. .11517  
/gene="SPAC186.03"  
/note="SPAC186.03, len:360, SIMILARITY:Schizosaccharomyces  
malidevorans., CAB69634, 1-asparaginase precursor., (356  
aa), fasta scores: opt: 1636, E():0, (70.2% identity in  
359 aa)"  
/codon\_start=1  
/label=SPAC186.03  
/product="l-asparaginase precursor"  
/protein\_id="CAB75867.1"  
/db\_xref="GI:7024420"  
/translation="MWRSIISFLFFSVALCQFLFQKRSSNISDFISFNASLPNVTIP  
AMGGTIAGYASSSTETVDYAAGSVGIATLVDVPAIKNFNIRGVQVTVNGSEELTPA  
NVNLTLQILAEVAKPDVHGIVVTHGTDLSLEETAMFLDMTNTTKPIVVVGAMRBPSTA  
ISADGPMNLNNAVVAASNRISGRGTMLLNDRIGSAFYTTKTNGNMLDTPFKSYEAGF  
LGMILDQRPFFYPATPTGKVHFDVSNSTTELPAYEILLYGYQGLPNLAKAAVDLGAK  
GLVLAGMGAASWTDPGNEVIDGLISNQSIIPVYSHRTMDGFSDDYYNGIPSYFQNPQK  
ARYMLMLSINAGYSIQNITDIFSLEY"  
10579. .11475  
/gene="SPAC186.03"  
/note="Match to PF00710 Asparaginase, Asparaginase Score  
335.65"  
complement(13275. .13334)  
/note="Duplicated region in c977 S. pombe chromosome 1:  
sub telomeric region"  
complement(13811. .14338)  
/gene="SPAC186.04c"  
complement(13811. .14338)  
/gene="SPAC186.04c"  
/note="SPAC186.04c, len:175,  
SIMILARITY:Schizosaccharomyces pombe., CAB69639, putative  
transmembrane channel protein., (598 aa), fasta scores:  
opt: 609, E():4e-32, (64.4% identity in 149 aa)"  
/codon\_start=1  
/label=SPAC186.04c  
/product="possible pseudogene, similar to N terminal of  
transmembrane channel"  
/protein\_id="CAB75868.1"  
/db\_xref="GI:7024421"  
/translation="MNTSSRIQLPSSNDAHVYDGRSNEPKASKRSYVNLTRQMRPKDA  
LKMNISSPNLKDLSKFADPPDAQQWLEGYLAKGLEDSKTPRSNFVDPLYEELNARRKP  
NKPVWSLSGPLPHVLGNSVVEKLEARSASSVNSRLNSRTNSSVSLKGMGSSSWKN  
KIKNAVSNVTDQSKR"  
complement(13838. .14256)  
/gene="SPAC186.04c"  
/note="Duplicated region in c977 S. pombe chromosome 1:  
sub telomeric region"  
complement(15587. .16375)  
/gene="SPAC186.05c"  
complement(15587. .16375)  
/gene="SPAC186.05c"  
/note="SPAC186.05c, len:262,  
SIMILARITY:Schizosaccharomyces pombe, YD68\_SCHPO,

hypothetical 31.3 kd protein c1798.08c in chromosome i., (287 aa), fasta scores: opt: 777, E():0, (50.8% identity in 240 aa)"  
/codon start=1  
/label=SPAC186.05c  
/product="hypothetical protein with PFUPF0016"  
/protein\_id="CAB75869.1"  
/db\_xref="GI:7024422"  
/translation="MINVESPLNLETTLSLQFSGVQNIQMSISMIIGCELGDKSF IVTALLAYQYGRASVFFGSLYALFMTSPAVLVGRAAPFLPKSITHILGTLFLIFG VKMLSEKQVRESQSLNEFDKVEKIIVNEEDMKKTLELGLPASNRSSSTLKDKFKF VFSMCKNLFESKKFSRAFIKAFALIFVSELGDRSQIATIVMSAKVKVLDVFIGVNIH HMLCTMVAVIVGRYISNKIEMVKVLPFGGIVFMIFGILYIPQGF"  
misc\_feature  
complement(15590..16291)  
/gene="SPAC186.05c"  
/note="Match to PF01169 UPF0016, Uncharacterized protein family UPF0016 Score 278.33"  
complement(16780..17089)  
/note="719 (-1) 44 349 Tf2-type LTR"  
17656..18210  
/gene="SPAC186.06"  
17656..18210  
/gene="SPAC186.06"  
/note="SPAC186.06"  
/note="SPAC186.06, len:184, SIMILARITY:Saccharomyces cerevisiae, YHI9\_YEAST, hypothetical 32.6 kd protein in dap2-slt2 intergenic region.hypothetical 32.6 kd protein in dap2-slt2 intergenic region., (294 aa), fasta scores: opt: 219, E():1.8e-07, (26.5% identity in 189 aa)"  
/codon start=1  
/label=SPAC186.06  
/product="hypothetical protein"  
/protein\_id="CAB75870.1"  
/db\_xref="GI:7024423"  
/translation="MSIKLGYLKPLPSYKTSISPEVIVDLEKFLNLCEGSCSQCNTP FCVDVGRNAILQLPNGADVILNPNFQAI FECCSKNSLTGVQIFGMVNDGTIELRSP CPVHGYNEDPANGSGAGSVGVFFALNPSIISSDFAHLLPNQKILGRNALIRVAIKL SANGLYDIHVGGGSKICISGTAEI"  
complement(18480..18715)  
/note="degraded LTR"  
complement(20852..21850)  
/gene="SPAC186.07c"

Query Match 27.3%; Score 265.4; DB 8; Length 30176;  
Best Local Similarity 56.4%; Pred. No. 3.9e-66;  
Matches 519; Conservative 0; Mismatches 396; Indels 6; Gaps 1;  
QY 42 CAAAACAGTAAAGTTGTGGTAGTGGGAGTGGGAAGTGTGGTCTGCACAGCGTATAC 101  
Db 25060 CTAGTCAATTAAAGTTGTGATTGTTGGTGTGGAAATGTGGTTCAACTACTGCTTTTAC 25001  
QY 102 GTTGCTTCTCAGCGGCATCGTTTCCGAGATTGTCCTTATCGACGTGAACAAAGCAAGC 161  
Db 25000 ACTACTACTAAGTGGATTAGCTGCTGAATCGTAATCATATTGACTTGAACAAAGCAAGC 24941  
QY 162 AGAGGTGAAGCATGGACTTAAACACGACGACCTTCAAATACAAAGTCTCGAG---- 217  
Db 24940 AGAAGGTGAAGCTATGGACTTGAATCATGACGCTCCACTTTCCCATGAGACTCGCGTATA 24881  
QY 218 --CGGTGATTATCTCGACTCGGCTGGCGGCCCATTTGTTATTGTACATGTGGGATTAA 275  
Db 24880 TTTGGGTGACTATAAAGATTGTAAGACGCAACGGCAGTTGTAATTACAGCTGGAAAGAA 24821  
QY 276 CCAAAAAAATGGACAAACAAGGATGGATCTTGCTGCAAAAAATGCCAACATTATGCTGGA 335  
Db 24820 TCAAAAGCCAGGTGAACAAAGAATGGATCTACTTAAAGCAAATATTTCCATCTTTAAGGA 24761  
QY 336 AATCATCCCCAATGTGCCAAATATGCTCTGATACCATCTGCTTATTGCCAGCAATCC 395  
Db 24760 AATCCTTAGAGAACTCACTAAATATACAAAGATGCCATCTCTTAGTTGCAACCAACC 24701  
QY 396 TGTCGATGTTTTCACCTATATTAGCTATAAGCGGCTCAGGGTTTCCACTAAGCAGAGTTAT 455  
Db 24700 TGTGGATGTACTTACTTACGCTACATTAAACTCACAGGATTTCCAGCTGAACGAGTCAT 24641

QY 456 CGGCTCAGGTACAGTTCTGGATACTGCTCGTCTTTTAAATACATCCTCGGAGAGCACTTCAA 515  
Db 24640 AGGATCCGGAACAATTATTGATACTGCAGGATTTCCAATATCTTATCGGAAAGTTATATGG 24581  
QY 516 GATCTCATCGGACAGCATCGATCGATCCCTGTGTAAATTGGAGAACATGGTGTGATTCGGGTGTGCC 575  
Db 24580 CCTTGATCCTCAAAGTGTAATGCCGATATCATTTGGAGAGCATGGGATTCGGAATTGGC 24521  
QY 576 TGTCTGTGCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTAATCGCGAAAAAGCCAA 635  
Db 24520 TGTATGTTCCACGCCAGTATTGCGAGACTTTTCACTTGCAGATTTCTGTGAAGAATCTGA 24461  
QY 636 CCACATATTGTGATCAGAATGCGTTCCATAGAATCTTTGAGCAACGCGAGACGCTGCTTA 695  
Db 24460 GACGAAGTATGATGAACAAGCAATTGAATGAATGCTTTAAAGAAAAACAAAAATGCCGCTTA 24401  
QY 696 CGATATCATCAAGCGCAAGGCTATATACTTCAATGGAATCGCAGCGGGATTACTTTCGCAT 755  
Db 24400 TGATATTATTCAAAGAAAAGGATCCACGGAATATGGCGTTGCTGCAGGTCTCGTTCGTAT 24341  
QY 756 AGTAAAGCGGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAACCGTTGGTGATTA 815  
Db 24340 TCTAGTGTCAATTATACGGGATGAAATGCACACTATTGACAGTATCTGGATTAGATCTTA 24281  
QY 816 TTTTGGGTTGAACAAATTGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGGCTCA 875  
Db 24280 CTCAACATTTGGTGATGTTTGTCTTAGTATGCCAAGAAAGCTAAACAAAGATGGTGTCTCA 24221  
QY 876 TCAAGTGGCTGAACTTTTCACTCGATGAGAAAGGAATAGAAATTGATGGAATAATCAGCTAG 935  
Db 24220 TCGATAAATTAATGCAAAGCTTTCCAAAAGACGAGGATGCCAAGCTTGTAGAATCCGTGAA 24161  
QY 936 TCAGATCAAATCAGTGATTGA 956  
Db 24160 AAGCATCAAGCATGCTATTGA 24140

RESULT 2  
BACLDHAA  
LOCUS BACLDHAA 1423 bp DNA linear BCT 26-APR-1993  
DEFINITION B.megaterium L-lactate dehydrogenase gene.  
ACCESSION M22305  
VERSION M22305.1 GI:143135  
KEYWORDS L-lactate dehydrogenase.  
SOURCE B.megaterium (strain DSM 090) DNA, clone pMH/pML.  
ORGANISM Bacillus megaterium  
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
REFERENCE 1 (bases 1 to 1423)  
AUTHORS Waldvogel,S., Weber,H. and Zuber,H.  
TITLE Structure and function of L-lactate dehydrogenases from thermophilic and mesophilic bacteria. VII. Nucleotide sequence of the lactate dehydrogenase gene from the mesophilic bacterium Bacillus megaterium. Preparation and properties of a hybrid lactate dehydrogenase comprising moieties of the B. megaterium and B. stearothermophilus enzymes  
JOURNAL Biol. Chem. Hoppe-Seyler 368 (10), 1391-1399 (1987)  
MEDLINE 88107005  
PUBMED 3122782  
REFERENCE 2 (bases 1 to 1423)  
AUTHORS Suter,F.  
JOURNAL Unpublished (1989)  
COMMENT [2] revises [1].  
Draft entry and computer readable copy of sequence [1] kindly submitted by S.Waldvogel 02-OCT-1987, and [2] by F.Suter 16-MAY-1989.  
Author address [2]  
F.Suter  
ETH-Hoenggerberg  
Inst. f. Molekularbiologie und Biophysik  
CH 8093 ZUERICH  
Switzerland.  
Location/Qualifiers  
FEATURES



```
source      1. .1423
/organism="Bacillus megaterium"
/db_xref="taxon:1404"
250_ .1206
/note="L-lactate dehydrogenase"
/codon_start=1
/transl_table=11
/protein_id="AAA22566.1"
/db_xref="GI:143136"
/translation="MKQFTPTKTRKVAVIGTFVGSYAFSMVNOQIANELVLIDMNR
EKAEGEARDINHGMPFATPMKIWAGDYKDCADADLAVITAGANQAPGETRLDLVEKNV
KIFECIVKIDIMNSGFDGIILVATNPVDILAHVTOKVSLPGRVIGSGTILDTARFRY
LLSDYFEVDSRNVHAYIMGEHGDTEFPVWSHAQIGGVKLEHFINTAIEKEPDMOHLF
EQTRDAAYHIINRKGATYYGIAMGLVRITKAILDDENSILTVSALLEGQYGISDVYIG
VPAIINKNGVRQIIELNLTPEHQOQLEHSASILKQTRDRAFY"
BASE COUNT 449 a 263 c 277 g 434 t
ORIGIN
Query Match      22.3%; Score 216.4; DB 1; Length 1423;
Best Local Similarity 54.9%; Pred. No. 6.3e-52;
Matches 496; Conservative 0; Mismatches 396; Indels 12; Gaps 3;
QY 41 CCAAAACAGTAAAGTTGTGGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATA 100
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 269 CAAAACACGAAAGTTGCCGTTATCGGAACCTGGTTTGTGGCTCAAGTACGCTTTT 328
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 101 CGTTGCTTCTCAGCGGCATCGTTTCCGAGATTGTCCTTATCGACGTGAACAAAGACAAG 160
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 329 CAATGGTGAATCAAGGTATTGCCAATGAATTAAGTTAATCGATATGAACAAAGAAAAG 388
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 161 CAGAGGGTGAAAGCATGGACTTAAACCACGACGACCTT-----CAATACAAAGTCTC 214
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 389 CAGAAAGTGAAGCACGCTGATATCAATCATGGAATGCCATTTGCCACACCGCATGAAATCT 448
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 215 GAGCGGGTGATTATCCTGACTGCGCTGGCGGGCCCATTTGTTATTGTACATGTGGGATTA 274
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 449 GGGCTGGAGATTATAAAGACTGTGCTGACGCTGAATTAGCAGTTATTACAGCGGGCGCTA 508
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 275 ACCAAAAAATGACAAACAAGGATGGATCTTGCTGCAAAAAAATGCCAACATTATGCTGG 334
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 509 ATCAAGCTCCAGGGGAAACACGCTTAGATCTAGTTGAAAAAACAACGTTAAATTTTCGAAT 568
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 335 AAATCATCCCCAATGTGCCAAATATGCTCCTGATACCATCCTGCTTATTGCCACGAATC 394
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 569 GCATTGTAAAGATATTATGAACACGGGATTTGACGGCATCATTTTATGGCAACAATC 628
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 395 CTGTCGATGTTTTGACCTATATTAGCTATAAGSCGTGAGGTTTCCACTAAGCAGAGTTA 454
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 629 CAGTTGATATTCTCGCACACGTTACACAAAAGTATCAGGATTACCAACGGACGGGTAA 688
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 455 TCGGCTCAGGTACAGTTCTGGATACTGCTCGTTTAAATACATCCTCGGAGACACTTCA 514
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 689 TTGGTTCAGGAACGATTTCTTGACACAGCTCGTCTCGCTACTTGTTAAGCGCACTATTTCG 748
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 515 AGATCTCATCGGACAGCATCGATGCCCTGTGTAATTGGAGAACATGGTGATCGGGTGTGC 574
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 749 AAGTAGATTCTCGCAACGCTCCACGCTTATATTATGGGGGAACATGGAGATACGGAATTC 808
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 575 CTGCTCGTCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAAGCCA 634
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 809 CTGTTTGGAGCCACGCGCAAAATTGGCGGTGTGAAGCTCGAACATTTTATCAATACTGCCG 868
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 635 ACCACATATTTGATCAGAATCGGTTCCATAGAACTTTGAGCAAAACCGGACGCTGCTT 694
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 869 CTATTGAAAAAGAACCGGATATGCAGCATCTATTTC--GAACAAACCCCGGATCGGGCTT 925
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 695 ACGATATCATCAAGCGCAAGGCTATACCTTCATATGGAATCGCAGCGGGGATTACTTCGCA 754
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 926 ACCATATTATTAAATCGAAAAGGACGCACTTATTACGGAATTGCAATGGGGCTTGACGCA 985
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 755 TAGTAAAGCGGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAAC---CGTTGGTG 811
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 986 TTACCAAGGCTATTTTAGATGATGAAAAATCTATTTTAAACAGTATCTGCTTTATTAGAAG 1045
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
QY 812 ATTATTTTGGGTTGAACAAATTGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGG 871
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1046 GACAATACGGTATTTCTGATGTATATCGCGGTACCGATCATTAATAAAACGGCG 1105
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 872 CTATCAAGTGGCTGAACCTTCACTCGATGAGAAGGAATAGAATTGATGGAAAAATCAG 931
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1106 TCGTCAAATATTGAATTGAATTAACTCCTCAGCAACAGCAGCAGCTCGAGCACTCTG 1165
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 932 CTAG 935
| | | |
Db 1166 CTAG 1169
| | | |
RESULT 3
AE006274/c
LOCUS
DEFINITION
Lactococcus lactis subsp. lactis IL1403 section 36 of 218 of the
complete genome.
ACCESSION
AE006274 AE005176
VERSION
AE006274.1 GI:12723239
SOURCE
Lactococcus lactis subsp. lactis.
ORGANISM
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE
1 (bases 1 to 10029)
AUTHORS
Bolotin,A., Wincker,P., Manger,S., Jaillon,O., Malarme,K.,
Weissenbach,J., Ehrlich,S.D. and Sorokin,A.
TITLE
The complete genome sequence of the lactic acid bacterium
Lactococcus lactis ssp. lactis IL1403
JOURNAL
Genome Res. 11 (5), 731-753 (2001)
MEDLINE
21235186
PUBMED
11337471
REFERENCE
2 (bases 1 to 10029)
AUTHORS
Bolotin,A., Wincker,P., Manger,S., Jaillon,O., Malarme,K.,
Weissenbach,J., Ehrlich,S.D. and Sorokin,A.
TITLE
Direct Submission
JOURNAL
Submitted (09-JAN-2001) INRA, Genetique Microbienne, Domaine de
Vilvert, Jouy en Josas 78352, France
FEATURES
Location/Qualifiers
1..10029
/organism="Lactococcus lactis subsp. lactis"
/strain="IL1403"
/sub_species="lactis"
/db_xref="taxon:1360"
122..412
/gene="ydhD"
/note="L173048"
122..412
/gene="ydhD"
/note="EVIDENCE BY HOMOLGY BIO14.04 Transposon related
functions"
/codon_start=1
/transl_table=11
/product="HYPOTHETICAL PROTEIN"
/protein_id="AAK04467.1"
/db_xref="GI:12723240"
/translation="MARRKFDQFKNSAVKLIIEGYSVKEVSQLELVHANSLYRWVQ
EVEEYGESAFPGNGTALADAQHKKLLEKENRYLKEELLELLKKFRVFLKRSK"
508..1269
/gene="tra904D"
/note="L0460"
508..1269
/gene="tra904D"
/note="EVIDENCE EXPERIMENTAL PMID:1328163 BIO14.04
Transposon related functions. BELONGS TO THE
IS3/IS150/IS904 FAMILY OF TRANSPOSASE. There are 9 such
elements in the chromosome"
/codon_start=1
/transl_table=11
/product="transposase of IS904D"
/protein_id="AAK04468.1"
```

/db\_xref="GI:12723241"
/translation="MHRPPSQQVEREILSEKIKAVFHEHKGRYGAVRITKVLHNTGI
MTNTRVGKLMHLMGLYAKGSRKYKYHNRKGASLRPNLNIQIFKATAPNKVWLGDM
TYIPKEGTLYLAVNIDVFSRKIVGWSISSRMQDKLVRCDFLQACGKEHPQGLIVHT
DQGSQYSSRYQSTLRQVGAQSSMSRKGNPYDNAMMESFYKTLKRELINDAHFETRAE
ATOEIFKYIETYNTKRMHSGLDYKSPKDFEKYS"
1326. .1883
/gene="ydhE"
/note="L37953"
1326. .1883
/gene="ydhE"
/note="EVIDENCE BY HOMOLOGY BIO14.04 Transposon related
functions"
/codon\_start=1
/transl\_table=11
/product="HYPOTHETICAL PROTEIN"
/protein\_id="AAK04469.1"
/db\_xref="GI:12723242"
/translation="MVKYSIELKQRIIQDYLSGKGGSTYLAKLHNVGSSSQVRHWIRN
YRAGLPTAHSKVNKNYSMELKENAVQCYLTDTLTVEAVARKFEITNFTLLASWNHF
KIYGEVPISKGRKKLESIASSMTNPNDQSRIKELEQELRYAQIEVAYLKGLRRL
EKNALMNKNQDSSTSVKPSNSKKS"
1880. .2698
/gene="tral077C"
/note="L0468"
1880. .2698
/gene="tral077C"
/note="EVIDENCE EXPERIMENTAL PMID:1328163 BIO14.04
Transposon related functions. BELONGS TO THE
IS3/IS150/IS904 FAMILY OF TRANSPOSASE. There are 7 such
elements in the chromosome"
/codon\_start=1
/transl\_table=11
/product="transposase of IS1077C"
/protein\_id="AAK04470.1"
/db\_xref="GI:12723243"
/translation="MKVTGPKATYYVWNCFERVNKDELIKEMLKIRQEHANAGYR
PMRELLKQRGYHVNHKVORLMMKGLRLVTSYWHKSRKNSYKGKVGTVAKNKLHRRF
RTSIPHQKITTDTEFKYEDGIQKCYLNPYIALFNSEVISYHISKQPSYQSIDIAL
NOALATSDCPYRRTFHSQGWYQMRDYVSKLKSRIHQSMRSKGNCHDNSVMENFF
GLLKQEIYGHIFSSFEELQVIVIRYNTKRIKQKLNWMSPIQFRLNYQNN"
complement(2772. .4256)
/gene="lyss"
/note="L0347"
complement(2772. .4256)
/gene="lyss"
/note="EVIDENCE BY HOMOLOGY BIO12.01 Amino acyl tRNA
synthetases"
/codon\_start=1
/transl\_table=11
/product="lysyl-tRNA synthetase (EC 6.1.1.6)"
/protein\_id="AAK04471.1"
/db\_xref="GI:12723244"
/translation="MAEIEELNDQMKVRREKMENLREAGIDPFGHKFTRTHNSOELHE
AYDEKTEELHELALSGIVAGRLMTRKGKGVFAHLQDREGQIQLYVRKDEVGSENY
EIFKKADLGDFLGVEGIMKTDMGELSIKAKKLTFLSKALRPLPEKFHGLTDTETRYR
KRYLDLISNKESFNRFVTRSKIISEIRRYMDGRGYLEVETPVLNNEAGGAAARPFYTH
HNSLDIDMALRIATELHLKRLJVGMEKVYELGRVFRNEGMDWTHNPEFTTMSYEAY
ADFEDIMDLTEGIFQHVAKTVVGDVLEYDGEINVGKPKRVHMVDAIKEVAGVDWF
PEMTFEATALAKEHDIHVEKHTSVGHIINEFFEKYVEETLIQPTFVFGHPKEISPL
AKMNEKDPRTDRFELFINGKEYANAFSELNDPIDQLEFFEAQAKAKELGDDEATGVD
YDYVEALEHGMPTGGIGIGIDRLVMLFTGTTSIDRVLLFPPTMK"
complement(4455. .5273)
/gene="rlrG"
/note="L0220"
complement(4455. .5273)
/gene="rlrG"
/note="EVIDENCE BY HOMOLOGY BIO09.04 LysR-family
regulators"
/codon\_start=1
/transl\_table=11
/product="LysR family transcriptional regulator"
/protein\_id="AAK04472.1"

/db\_xref="GI:12723245"
/translation="MFALLETFIAVFETKSFTRAANQCFISQPTATVRIKKLEELKV
HLFSRGQHQEVIPTESAHLLYPKALKTLNDWNELOQFSLKNQTPSRKPFKIAAHSASAT
TVLPLIFKNLIPLEQNLIELMLLDSQEVFDLIRNHEIHFGIIEKPMMDETIETFFLF
KDELVLAGNPESDIFFIIEAGSGVAHYTHNYLKENNLNPKNLVSVNNNDMIVSQLKAG
IGLSLISKRFINENIAFQNLGERYQRIFYGLSFLSEQDKLLKTVIKKIKNLSEF"
5437. .6432
/gene="ydhF"
/note="L178384"
5437. .6432
/gene="ydhF"
/note="EVIDENCE BY HOMOLOGY BIO15.01 HYPOTHETICAL"
/codon\_start=1
/transl\_table=11
/product="HYPOTHETICAL PROTEIN"
/protein\_id="AAK04473.1"
/db\_xref="GI:12723246"
/translation="MSILPGFLLSFAIAIVSYLNRFFIHSLSGATIAILLGIILGNL
YFKQVTLFACTAWSKKLLEFSVMFLGATVTFQTIKLGSGVGFILIQMISTIIFVL
FMGKKLGFSANVSALMASGNVACGSSAIAAVEPVIGAETSEKRTSIAMVNLMTILML
SLPFLGTWFMGNDDLRGALIGGTEQSVQGVASATVNPNTTTLATLKFIMRIIMLV
FVVLXFGFRSKKQKNEQGPTQIKIKRNSFLPWVGLVLCITLDTLIHFVPEVSATA
KFLSGWCETIALAAIGRLNLVEFIKAGKKLLIYGLSTLVFQVVLALILISLLIK"
complement(6645. .7589)
/gene="ldhB"
/note="L0018"
complement(6645. .7589)
/gene="ldhB"
/note="EVIDENCE BY HOMOLOGY BIO06.09 Glycolysis. can be
malate dehydrogenase"
/codon\_start=1
/transl\_table=11
/product="L-lactate dehydrogenase (EC 1.1.1.27)"
/protein\_id="AAK04474.1"
/db\_xref="GI:12723247"
/translation="MKITSRKVVVIGTGVGTSIAYSMINQGLVNELVLIDVNQDKAE
GEALDLLDGVSWGQENIVRAGDYKDCKNADIVVVTAGVNQKPGQSRLLDLVNTNAKIM
RSIVTQVMDSGFDGIFVIASNPVDILTYVAVETSGLDQSRIVGTGTTLDTTRFRKELA
TKLEIDPRSVHGYYIGEHDSEVAVWSHTTVGGKPILEFIVNKKIGIVEDLSNLSKV
KNAAYEIIDKKQATYYIGIMSTARIKAILNNEQAILPVSAYLARGEYQGEVFTGVPS
IVNQGVREIIELNIDAYEKKQFEKSVSQLKEVIESIK"
complement(7883. .8704)
/gene="rlrD"
/note="L0217"
complement(7883. .8704)
/gene="rlrD"
/note="EVIDENCE BY HOMOLOGY BIO09.04 LysR-family
regulators"
/codon\_start=1
/transl\_table=11
/product="LysR family transcriptional regulator"
/protein\_id="AAK04475.1"
/db\_xref="GI:12723248"
/translation="WPSLFEFIVVYETKSFTLACKYLFISQPTVTVRIKKLEELKS
TLFLRGKHQEIIPTEAAILFYPKAIAYLKKWEEDAQVQKSLAKHPFKIGVSHSAAAL
SIMPQIFKVFEGELEHLDVEINMYDSEKVFELVANHDLHFGIIERLLISDQETETFFLF
LDELVLAGNTDSEFTFTRESGSGIGYYIKRYLKSAPQARNIVRMNSNEMIISHIKAG
LGSSLSIKSFLTDDIPYQELGSNYHREFLGVSFDEEKDPMIQKLISKIKKETEIH"
8859. .9833
/gene="ydia"
/note="L181867"
8859. .9833
/gene="ydia"
/note="EVIDENCE BY HOMOLOGY BIO13.01 TRANSPORT AND BINDING
PROTEINS. General. tellurite resistance protein or related
permease, evidenced by COGnitor"
/codon\_start=1
/transl\_table=11
/product="permease"
/protein\_id="AAK04476.1"

Matches		507;	Conservative	0;	Mismatches	428;	Indels	15;	Gaps	3;
QY	29	TAAGAACTGATGCCAAACAGTAAAGTTGTGTAGTGGAGTGGAGTGGTGGTCTG 88								
Db	7594	TAAAATGAAAATTACAAGCAGAAAAGTAGTTTATAGGAACAGGTTTGTGGACAA 7535								
QY	89	CCACAGCGTATACGTTGCTTCTCAGCGGCATCGTTCCGAGATTGCTCCTTATCGACGTGA 148								
Db	7534	GTATTGCTTATTCATGATTAAATCAAGGACITGTTAATGAGTTGGTTTGATAGATGTTA 7475								
QY	149	ACAAAGACAAAGCAGAGGTGAAAGCATGGACTTAACACCAAGCAGACCTTC----- 200								
Db	7474	ATCAGGATAAAGCTGAGGGTGAAGCCTTAGATTATTAGATGGCGTTTCTTGGGGCAAG 7415								
QY	201	-AAATACAAGGCTCGAGCGGTGATTATCCTGACTGCGTGGCGGCGCATTTGTTATTG 259								
Db	7414	AAAATGTTATTGTGAGAGCTGGAGATTATAAGGATTGTAAAAACGCTGATATTGTGTCG 7355								
QY	260	TCACATGTGGATTAAACCAAAAAAATGGACAAACAGGATGGATCTTGTGCAAAAAATG 319								
Db	7354	TTACGGCTGGGTAATAACAAAACCAAGTCAATCAGCTTTAGATTGTTGTTAATACAAATG 7295								
QY	320	CCAACATTATGCTGGAATCATCCCCAATGTTGCCAAATATGCTCCTGATACCATCCTGC 379								
Db	7294	CTAAAATTATCGCTCAATCGTGACTCAGGTTATGGACTCAGGTTTGTGATGTTTGTG 7235								
QY	380	TTATTGCCACGAATCCTGTGATGTTTTGACCTATATTAGCTATAAGCGTCAGGGTTTC 439								
Db	7234	TAATTGCTTCAAAATCCTGTGGATATTCTTACCCTATGTCGCTTGGGAACTTCTGGTCTTG 7175								
QY	440	CACTAAGCAGAGTTATCGGCTCAGGTACAGTTCTGGATACTGCTCGTTTTAAATACATCC 499								
Db	7174	ATCAATCTCGAATTGTCGGAACCTGGTACAACTTTGGATACTACTCGTTTTAGAAAAAAGAT 7115								
QY	500	TCGGAGAGCACTTCAAGATCTCATCGGACAGCATCGATGCCTGTGTAATTGGAGAACATG 559								
Db	7114	TGGCCACAAAGCTAGAAATTGATCCACGGAGCGTTCATGGATATATTATTGGCGAACATG 7055								
QY	560	GTGATTCGGGTGTGCTGTGCTCTCTTACCAACATCGACGGCATGAAGTCCGGGATT 619								
Db	7054	CGGATTCGAAGTTGCAAGTTGGTGCACACACGCGGTAGGTGGAAACCTATTCTTGAAT 6995								
QY	620	ACTCGAAAAAGCCACACATATTTGATCAGAAATGCGTTCATAGAAATCTTTGAGCAA 679								
Db	6994	TTATTGTTAAAAATAAAAAATAGGGTTGAAGTTTGTCTAATTTGTCAAATAAGGTAA 6935								
QY	680	CGCGAGACGCTGCTTACGATATCATCAAGCGCAAGGCTATACTTTCATATGGAATCGCAG 739								
Db	6934	---AAATGCTGCCTATGAGATTATTGTATAAAAAACAGGCTACTTATTACGCGATTGAA 6878								
QY	740	CGGGATTACTTCGATAGTAAGCGGATTTTAGAGGATACAGGATCCACACTTACAGTTT 799								
Db	6877	TGAGTACAGCTAGAAATTGTCAAGGCCATCTTAAATAATGAACAAGCAATTCTACCAGTTT 6818								
QY	800	CAACC---GTTGGTGAFTATTTTGGGGTTGAACAAATTGCTATAAGCGTCCCTACCAAC 856								
Db	6817	CTGCCATTATTACGTGGTGAATATGGTCAAGAAGGTGTCTTTACCGGTGTTCATCTATTG 6758								
QY	857	TCAATAAAAGTGGGCTCATCAAGTGGCTGAACCTTTCACCTCGATGAGAGGAAATAGAAT 916								
Db	6757	TAAATCAAAATGGTGTAGAGAAATCATTTGAACCTAAAATATTGATGCTTTATGAAAAGAAC 6698								
QY	917	TGATGAAAAAATCAGCTAGTCAGATCAAAATCAGTGAATTGAGCATCTGGAG 966								
Db	6697	AATTGAAAAAATCTGTCAGTCAACTTAAAGAGGTAATAGAAATCTATTAAAG 6648								

RESULT 4  
AE007540  
LOCUS  
DEFINITION  
ACCESSION

11158 bp DNA linear BCT 27-JUL-2001

AE007540  
Clostridium acetobutylicum ATCC824 section 28 of 356 of the complete genome.  
AE007540 AE001437

VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
gene  
CDS  
gene  
CDS  
gene  
CDS  
gene  
CDS

AE007540.1 GI:15023104  
Clostridium acetobutylicum.  
Clostridium acetobutylicum  
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.  
1 (bases 1 to 11158)  
Nolling,J., Breton,G., Omelchenko,M.V., Markarova,K.S., Zeng,Q., Gibson,R., Lee,H.M., Dubois,J., Qiu,D., Hitti,J., Wolf,Y.I., Tatusov,R.L., Sabathe,F., Doucette-Stamm,L., Soucaille,P., Daly,M.J., Bennett,G.N., Koonin,E.V. and Smith,D.R.  
Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum  
J. Bacteriol. 183 (16), 4823-4838 (2001)  
21359325  
11466286  
2 (bases 1 to 11158)  
Childress,D., Zeng,Q. and Smith,D.R.  
Direct Submission  
Submitted (24-JUL-2001) GTC Sequencing Center Production, Finishing, and Bioinformatics teams, Genome Therapeutics Corp., 100 Beaver Street, Waltham, MA 02453-8443, USA  
Location/Qualifiers  
1. .11158  
/organism="Clostridium acetobutylicum"  
/strain="ATCC 824"  
/db\_xref="ATCC:824"  
/db\_xref="taxon:1488"  
122. .1063  
/gene="CAC0267"  
122. .1063  
/gene="CAC0267"  
/codon\_start=1  
/transl\_table=11  
/product="L-lactate dehydrogenase"  
/protein\_id="AAK78248.1"  
/db\_xref="GI:15023105"  
/translation="MKNTKISVIGAGFVGSSVTFALMNGGLASEIIVDVNKKABG EAMDLSHGAAFVKPVVKSVDYKDTGSDIIVITAGAAQKPGETRLLELNKNYNIPKS IPEVVKYNPNAILLVVSNDPVDILTYITYKLSGPKSRVIGSTVLDTSRFRYMLSEH FEIDVRNIHTYIMGEHGDSEIATWSLTNIAGMDVNECEASCKKCDGSLKYKIYDDVK NAAHVIEKGYAYVALAVKRIVEAILRDENSILTVSSLLLEGQYGIKDVYMGVPSI VGINGVKDIIIEVPLNDEKNLTDsAKTLKESLDSIP"  
1134. .1910  
/gene="CAC0268"  
1134. .1910  
/gene="CAC0268"  
/codon\_start=1  
/transl\_table=11  
/product="ABC transporter ATP-binding protein"  
/protein\_id="AAK78249.1"  
/db\_xref="GI:15023106"  
/translation="MLYIYNKEIGGVIVIELNVNKSNGKNKAVDNLNFTINSGEI FGFLGPNAGAKSTTIKMTIGIIKADSGEIKVDGLNKENPIEAKRKIGYVPDSDPMFL RLKGIENLNFMAIDYDVTSDRKERIESLSKYFDMESALDDKIQSYSHGMQRQKIVVMG ALIHPDQWILDEPMTGLDPKASYNLKEMMRKHADSGKTFFSTHLEVAEKLCDRIG IVSKGKIVFCGTWNEKQKLQEEGSSLERMFLELVEDGEK"  
1867. .3543  
/gene="CAC0269"  
1867. .3543  
/gene="CAC0269"  
/codon\_start=1  
/transl\_table=11  
/product="Uncharacterized membrane protein"  
/protein\_id="AAK78250.1"  
/db\_xref="GI:15023107"  
/translation="MKECFWSWLKMEKNKFLILTKYLLLGSGYSGRKKKGIFSKSTV ANLVKFIILIFGATAGSIDLAAYMGLKIVHMEYVLLQMNYSAITAIIFFGIFSTM GVYFNSNDINFLIPMLKPEITIVASKFTVSLIOEYFTIVILLPCTLVYGIASGMGVI FVILSILFMLEFEPILPLAAALLISIFIMPLINKSKKDLFKTLGGLAVGFLGINIL TRMMGGSSSAAELVNKVRKNSMTMLNVPLGINSKASKLVNSNLLYTMFLILINACSFVA IYVLAKNLYFKVGQLSESSSKRKIDSKMKNKTQKKSILFSYTMKELKILFRTPAY FVNCVMSSFIFFVVILIPMLFNTSDKKGLSSALSINGPQGFRIILVTASVVFGLIMGT



```

MNSVTSTSIREGSNFFVMKYIPIKYSQIAAKIFSGIFVGAISLILIDIAAFFVVKI
PVYFAYMIITGILGIVFTSIIGMYIDLSPKLEWDTEQKAVKSNINSFINFLVILP
AAIIIVICIFINVNWNIFFGILSLVLLALDYVTKVVCISGEKKLDRIES"
complement (3604. .4341)
/gene="CAC0270"
complement (3604. .4341)
/gene="CAC0270"
/codon start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="AAK78251.1"
/db_xref="GI:15023108"
/translation="MISIIIPDNQIKKINLYWQNIIFYKPSDKDTESTQIIPDNLKNSG
KISIVHNDKSFIKIDPCILSKISDFVNSHPYKTLYSKDFEDYFKNNIRRYENAYPC
YLTEPSYSDVLNSENSEYVRKLLPSDTKVLNFKKSHSEIHKTTLIENYPILLGCFY
KDSLVSASLDYFKNDIANLNIITHPSFRDNTIISKTLIAELCQFCLKEGKILQYNFL
DKDERAKKLAQSMGFNLVYFKENLVLK"
4598. .5014
/gene="CAC0271"
4598. .5014
/gene="CAC0271"
/codon start=1
/transl_table=11
/product="Predicted esterase"
/protein_id="AAK78252.1"
/db_xref="GI:15023109"
/translation="MNKAITRIKVRYAETDKMGIVYHANYVYFEAAREDLIEGAGIK
YSDMEDIGIMMPLVETCKYHEGAKYGDYILVETTLGKLSPIKVEINRYVLRESDGKL
LAEGQTTQVFVDAQNFKIIKLMKSYPEWEKLOKIK"
5157. .7019
/gene="CAC0272"
5157. .7019
/gene="CAC0272"
/codon start=1
/transl_table=11
/product="Amino acid transporter"
/protein_id="AAK78253.1"
/db_xref="GI:15023110"
/translation="MVNKFDFDLICRPLANEQSSSEKYNVPFGLAVMASDAISSVAYA
SQEILLVLVPIIGAASYAWLKGASLMIIGLLILTVSYVQIIKAYPQGGGAYIVAKEN
LGVPGLVAGASLLDYILTVAVSASSGVAAITSAFSLVPHSVLLAVVLIILITILN
LRGVSESAKMFSIPTYLFIFSMIFMIIFGIAKYIIYGAPSMHMIHQTLKPMGDVSIPL
ILKAFSSGCSALTGLEAVSNSVPNKEPQGKNATIMVLLTLCIFVIFGSSLLARFY
GAIPNDNVTVLAQIAYGVFSGKTFMFYVVOFTTAIILMACNTSFTGPFMLMSVIARD
GYAPRSFATRGRLSFGSIVFLSIIAGILVIIIFKADTHLLIPLYSVGVLFSFTLAQT
GMVIHWKNTNEKGKNTKRKLINGFGAAITLLTTGVIIVYKFLSGAWVIIILVPIVLGF
LH1KHYNVVARGLRADEEYVKNGCIGQTYNHLIIIVPTASLNKATLGALKYARSITS
VIALNVPSPNKEYMDLKHKWEELNTDILLVSKSPYRAIVPTLIEYIGVISKAAVPDE
KITVVLPPQFITRDKSGQILHNHTSFLREALLRYDNIIVSTYPFHLSPKODNEE"
7464. .9140
/gene="CAC0273"
7464. .9140
/gene="CAC0273"
/codon start=1
/transl_table=11
/product="2-isopropylmalate synthase"
/protein_id="AAK78254.1"
/db_xref="GI:15023111"
/translation="MLYNKYVKYPKVKLDNREWPDQIEKAPIWCSVDLRDGNQALPK
PMNVDEKIKMFKMLVDIGFKEIEIGFPSASETEYEFTRKLIEDKLIPEDTVITQVLTQA
REHLVKKTFEALKGVKTAIVHVYNSTSELQRRVVFKKDKDEVKSLAIKGAQMWKYISE
YTEYSESKFVFEYSPESFTGTGLDYALEVCEAVLNWVKPSKENKAIINLPSTVEMATP
NIYADQIEWFCKKLLNRESVILSLHTHNDRGTCASSSELGILAGADRVEGTLFGNGER
TGNLIDIMMALNMYSQGVDPLEFSLNNDIVEAYEECTKMVVHERHPYAGKLVPTAFS
GSHODAIRKGLKSLKEGNHWEVPYLAVDPHDLGVEYBEIIRINSQSGKGGTAYIME
SDFGFIILPKAMHPEFGKVIKKSDDELDCELSPEQIFKFPKBEYLENRSPLYLKNYKIH
SIONTEEEKNTVDIEAVISVNGKDTSEIGVNGPVDAPFNAMNNKYNKGCKFISYDEH
ALNIGSHSKAVAYVQIESQDKKYFGVGISDNIDTASINAIVSALNRSKLK"
9655. .11052
/gene="CAC0274"
9655. .11052
/gene="ansB"
9655. .11052
/gene="CAC0274"

```

```

/codon start=1
/transl_table=11
/product="Aspartate ammonia-lyase (aspartase) gene
ansB(aspA)"
/protein_id="AAK78255.1"
/db_xref="GI:15023112"
/translation="MDFRIESDSIGDKKVPKEVYVQTLRGADNFKITGLRLNEEFI
KSVVQIKKAAAITNLEIGELEVNIKAIKACDEVIAKFNEFDIIDPIQGGAGTSIN
MNAVEIANRAIEILGYEKGNYRIVHPNDHVNMGOSTNDVIPTAGKITALRLVINKLD
KLKELSNAFSEKAEQFDVKIMGRTOQLQDAVITLGOEFSAYSAVVKRDIARIQKAQD
ELKVVNIGGTAVGTGINADSRVVDRIVPRLSFVSLNKLKADDLIDGTQNLDSFASLS
ATIKTCALNLSKIANDLRLMSSGPRITGLGEINLPKQNGSSIMPGKVPNPIPEVINQI
AFNIAGNDVAISMAVEAQLELNAFEPVIFYNIFESIQTLTNGVDVTFIKNCILDITAN
EEICSEYVNNVSGIVTALCPYIGYAESAKIAKMAIKSGESIKKVVLREGVLKEKELNS
IINPLKMTKPGIVGK"
BASE COUNT      4051 a 1350 c 2112 g 3645 t
ORIGIN

Query Match      20.7%; Score 201.2; DB 1; Length 11158;
Best Local Similarity 53.3%; Pred. No. 2.2e-47;
Matches 475; Conservative 0; Mismatches 408; Indels 9; Gaps 2;

QY 52 AAAGTTGTGGTAGTGGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC 111
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 AAAATTTCAGTAATCGGAGCAGGATTTGTTGGTTCTTCTACAGTATTTCGCTAATGAAC 196

QY 112 AGCGGCATCGTTTCCGAGATTGTCTTATCGACGTGAACAAAGACAAAGCAGAGGTGAA 171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 197 GGTGGACTTGCATCTGAGATAGTTATAGTAGATGTAATAAAGATAAGGCAGAAGGGAA 256

QY 172 AGCATGGACTTAAACCCAGCAGCACCTT-----CAAATACAAGGTCTCGAGCGGTGAT 225
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 GCAATGGACTTATCTCATGTGTGCAGCTTTTGTAAACCTGTAGTTGTGAAATCTGGAGAT 316

QY 226 TATCCTGACTGCGCTGGCGGCCCATTTGTTATTGTACATGTGGGATTAAACCAAAAAAT 285
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 TATAAGATACAGAAGGTTTCAGACATAGTTATAATAACAGCGGAGCAGCACAGAAACCA 376

QY 286 GGACAAAACAAGGATGGATCTTGCTGTCAAAAAAATGCCAACATTATGCTGGAATCATCCC 345
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 GGTGAGACAAGACTTGAATTGATAAATAAGAACTACATAATATATTAAAGTCTATAGTTCT 436

QY 346 AATGTTGCCAAAATATGCTCCTGATACCATCCTGCTTATTGCCACGAATCCTGTGATGTT 405
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 437 GAA GTTGTAATAATATAATCCAAACGCAATTTTATTAGTAGTTTCTTAATCCAGTAGACATA 496

QY 406 TTGACCTATATTAGCTATAAGGCGTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCAGGT 465
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 497 TTAACCTATATTACCTATAAACTTTTCAGGTTTTTCCAAAGCTTAGAGTAATAGGTTCCGGT 556

QY 466 ACAGTTCTGGATACTGCTCGTTTAAATACATCCTCGGAGAGCAGCACTTCAAGATCTCATCG 525
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 557 ACAGTTCTTGATACATCAAGATTTAGATACATGTTAAGTGAACATTTTGAAATTGATGTT 616

QY 526 GACAGCATCGATGCTGTGTAATTGGAGAACATGGTGATTTCGGGTGTGCCCTGTCTGGTCT 585
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 617 AGAAATATACACACTTACATTTATGGGAGAACACCGGAGATTTCAGAAATAGCTACATGGAGT 676

QY 586 CTTACCAACATCGACGGCATGAAGCTCCGGGATTACTCGCAAAAAAGCAACACATATTT 645
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 TTAACAAAATATAGCAGGGATGCGATGTAATGAATATTGTGAAGCTTCTGCAAAAAATGT 736

QY 646 GATCAGAAATGCGTTCCATAGAAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATATCAT 705
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737 GATGATCATTTAAAAATATAAAATATATGACGATGTAAAAAATGCGGCGTATCATGTTATA 796

QY 706 AAGCGCAAAGGCTATACITTCATATGGAATCGCAGCGGGGATTACTTCGCATAGTAAAGCG 765
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 797 GAGAAAAAGGGTGCAACTTATTACGCTGTTGCACCTAGCTGTAAAAAAGAAATAGTAGAGCT 856

QY 766 ATTTTAGAGGATACAGGATCCACCTTACAGTTTCAACCGTTGGTGA--TTATTTTGGG 822
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 857 ATATTAGAGAGATGAAAAATTCAAATTTTGACAGTATCTTCTCTTCTTGAAGGTCAATATGGA 916

```







TAGIIMKSGIRAAAYETGKRIVVRAKAEIEENGRHKIIIVTELPYQVNKAKLIEVIA  
DLVKDKITGISDLRDESDREGMRMVIELKRDANPNVTLNLLYKHKMQDTFGVIMLA  
LVDNQPOLNLKQVLVHYINFQKDVITRRTOFELNKAKEHILEGLIIALDNIDEVI  
NIIRSKTSEIAKNTLIERFSLSEKQAAAILLEMRRLTGLERDKIENEYNELMKLID  
YLNAILASEERLLEVIKEELLEIKAKYSDERRTOIEKQENEDIEDLIOEEDVVITLT  
HTGIYKRISSDVYSAQKRGKGIGQAMTTEKDDFVENVLVTSHTSNLLFFTTKGKVVYKI  
KAYEVPDAGRTAKGTLNVLNIPLEADEKIQTVLSFKDVQEDGYLFMGTKLGIVKKTPL  
KDFKNIRKNGLIALNREGDELLRAKITYGDAIIIFVTDGNAIRFNETDVRPMDGRTA  
SGVKALSIREGDVAVCMDIAVEDEKLLVSENGYGKRTNLDYKVQNRGGIGLIITYKV  
SDTKGVVGATVCKEDDELMHYK"  
9369. .9620  
/gene="CPE0008"  
9369. .9620  
/gene="CPE0008"  
/note="83 aa, no significant homology."  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAB79714.1"  
/db\_xref="GI:18143665"  
/translation="MSLCIINSQVAIRINVSDisVTSRSAMGVKLMRTLEDEAVTII  
AKISGDNTNEEEEEITLLDESEDNTIIVEEVATEVKDEEN"  
10173. .11680  
/gene="rRNA-16S"  
10173. .11680  
/gene="rRNA-16S"  
11869. .14773  
/gene="rRNA-23S"  
11869. .14773  
/gene="rRNA-23S"  
14868. .14982  
/gene="rRNA-5S"  
14868. .14982  
/gene="rRNA-5S"  
14992. .15065  
/gene="trnA-Met"  
14992. .15065

gene  
rRNA  
gene  
rRNA  
gene  
rRNA  
gene  
tRNA

Query Match 19.1%; Score 185.8; DB 1; Length 301450;  
Best Local Similarity 52.1%; Pred. No. 9.6e-43;  
Matches 468; Conservative 0; Mismatches 422; Indels 9; Gaps 2;

QY 43 AAAACAGTAAAGTTGTGGTAGTGGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACG 102  
Db 142778 AAAACTAATAAATCTCAATATAGGTGCAGGATTGTGGTTCAACACAGCCTTCGCA 142719

QY 103 TTGCTTCTCAGCGGCATCGTTTCCGAGATTGTCCTTATCGACGTGAACAAAGACAAAGCA 162  
Db 142718 TTAATGCAAGATGGATTAGCATCCGAAATAGTAAATGTTGTATATAAATAAGGATAAGCA 142659

QY 163 GAGGGTGAAAGCATGGACTTAAACCACGACGACCTTCAAATACAAGTCTCG-----A 216  
Db 142658 CATGCTGAAGCTATGACCTAGCACAAAGGTGCGGCATTTGTTAAATCTGTGTATATAAAA 142599

QY 217 GCGGGTGATTATCCTGACTGCGCTGGCGGCGCATTTGTTATTGTACATGTGGGATTAAC 276  
Db 142598 TCAGGAGACTATGCAGATACAAAAGACTCTGCACATAGTTATAATAACTGCTGGAGTTGGT 142539

QY 277 CAAAAAATGGACAAACAAGGATGGATCTTGTCTCAAAAAAATGCCAACATTTATGCTGAA 336  
Db 142538 CCAAAGCCAGGTGAACAAGATTAGATATTATAATAAATAAACTTAAAAATCTTCCAAAGT 142479

QY 337 ATCATCCCCAATGTTGCCAAATATGCTCCTGTATACCATCCTGCTTTATTTGCCACGAATCCT 396  
Db 142478 ATAGTACCTGAGGTTGTTAATATATAGCCCAAACCTCAATACTTTTAGTAGTATCAAAACCA 142419

QY 397 GTCGATGTTTTGACCTATATTAGCTATAAGGCGTCAGGGTTTCCACTAAGCAGAGTTATC 456  
Db 142418 GTTGATATTTTAACTTTATATAACTTACAAGCTTTCTGGATTCCCTTAAAGAACGTGTTATT 142359

QY 457 GGCTCAGGTACAGTTCTGGATACTGCTCGTTTTTAAATACATCCTCGGAGAGCATTCAAG 516  
Db 142358 GGATCTGGAACCTGCTTAGATACTTCAAGATTAAATATATATGCTTTTCAGAACATTTTGAT 142299

QY 517 ATCTCATCGGACAGCATCGATCGCTGTGTAATTGGAGAAACATGGTGATTCGGGTGTGCCT 576  
Db 142298 ATAGACGCTAGAAATGTTCCACATATATAATTGGAGAACATGGAGACTCAGAAATAACA 142239

QY 577 GTCTGTCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTAATCGGAAAAAGCCAAC 636  
Db 142238 GCCTGGAGCTTAACTAATATTGCTGGAGCAAAATGTTGAAGAATATTGTAAACAGTTTGC 142179

QY 637 CACATATTGTGATCAGAAATCGCTTCCATAGAAATCTTTGAGCAAACGCGAGACGCTGCTTAC 696  
Db 142178 GCTAACTGCGATGGCTCCTTTAAAAAGGAACTTCTTGAAAAAGTTAAAAATGCAGCTTAT 142119

QY 697 GATATCATCAAGCGCAAAGGCTATATCTTATGGAATCGCAGCGGGATTACTTCGCATA 756  
Db 142118 GAAATAATAATAGTAAAGGCTATACAAACTATGCTGTAGCTTTGCCGTTACAAGAATA 142059

QY 757 GTAAAGCGGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAAC---CGTTGGTGAT 813  
Db 142058 GTTGAAGCAATCCTTCGTGATGAAAAATGCTATTTTAACTGTTTCTTCTCTATTGGAAGC 141999

QY 814 TATTTTGGGGTTGAACAAATTTGCTATAGAAGCGTCCCTACCAAACTCAATAAAAGTGGGCT 873  
Db 141998 CAATACGGAATTGATAATGTTTATTATTAGCAATGCCTACTATATAGTTGATAGAAGTGGAGCT 141939

QY 874 CATCAAGTGGCTGAACCTTTCACTCGATGAGAAGGAAATAGAAATGATGGAAAAATCAGC 932  
Db 141938 AGACAAATCTAGATGTCCCTATATCAAAATGAAGAAAAAGAAACCTTAATAAAATCCGC 141880

RESULT 7  
AX142689  
LOCUS AX142689 969 bp DNA linear PAT 31-MAY-2001  
DEFINITION Sequence 1411 from Patent WO0134809.  
ACCESSION AX142689  
VERSION AX142689.1 GI:14282216  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1 (bases 1 to 969)  
AUTHORS Kimmerly,W.J.  
TITLE Staphylococcus epidermidis nucleic acids and proteins  
JOURNAL Patent: WO 0134809-A 1411 17-MAY-2001;  
GLAXO GROUP LIMITED (GB)  
FEATURES  
source 1. .969  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="synthetic nucleic acid sequence"  
BASE COUNT 339 a 130 c 194 g 306 t  
ORIGIN

Query Match 18.9%; Score 183.8; DB 6; Length 969;  
Best Local Similarity 53.0%; Pred. No. 2.1e-42;  
Matches 473; Conservative 0; Mismatches 402; Indels 18; Gaps 3;

QY 51 AAAAGTTGTGGTAGTGGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCT 110  
Db 36 AAAAGTTGTTTAGTAGGAGACGGTTCGTTAGGTTTCAAGTTATGCAATTTGCTATGTTGAC 95

QY 111 CAGCGGCATCGTTTCCGAGATTGTCCTTATCGACGTGAACAAAGACAAAGCAGAGGTGA 170  
Db 96 TCAAGGAATTGCAGATGAATTTGTAATTATTGATATTGCAAAAGATAAAGTGAAGCAGA 155

QY 171 AAGCATGGACTTAAACCACGACACCTTCAAATACAAGGTCCTCG-----AGCGGG 221  
Db 156 CGTTAAAGATTTAAACCATGTTGCACCTTACAGTTCTTCCACCACTGACTGTAAAAGCTGG 215

QY 222 TGATTATCCTGACTGCGCTGGCGCGGCCATTGTTATTGTACATGTGGGATTAAACCAAAA 281  
Db 216 AGAATATGAAGATTGTAAAGATGCAGATTTAGTTGTTATTATACAGCAGGTGCACCTCAAAA 275









QY 462 AGGTACAGTTCTGGATACTGCTCGTGTGTTTAAATACATCCTCGGAGACGACTTCAAGATCTC 521  
|||||  
Db 1372 TGGTACAGTCTTGATAGTCAAGATTTCAGATATTTAATAAGTAAGAATTAGGTGTTAC 1313  
  
QY 522 ATCGGACAGATCGATGCTGTGTGTAATTGGAGAACATGGTATTCGGGTGTCCTGCTG 581  
|||||  
Db 1312 ATCAAGTAGTGTACGCTAGCATTTATAGGTGAACATGGTGACTCTGAACCTTGCAGTTTG 1253  
  
QY 582 GTCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCAAACCACAT 641  
|||||  
Db 1252 GTCTCAAGCAACGCTGGAGGTATTTTCAGTGTATGATACATTTGAAAGAAGAACTGGTAG 1193  
  
QY 642 ATTTGATCAGAAATCGTTCATAGAAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATAT 701  
|||||  
Db 1192 CGATGCTAA-----AGCGAATGAAATTTATATTAATACAAGAGATGCTGCTTACGATAT 1139  
  
QY 702 CATCAAGCGCAAGGCTATACTTCATATGGAATCGACGGGATTACTTCGCATAGTAAA 761  
|||||  
Db 1138 CATTCAGCTAAAGGATCTACGTATTATGGTATAGCTCTAGCACTATTACGTATTTCTAA 1079  
  
QY 762 GCGGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAA---CCGTTGGTGATTATTT 818  
|||||  
Db 1078 AGCTTTACTAAATAATGAAATAGTATTTGACAGTTTCTAGTCAACTTAATGGTCAATA 1019  
  
QY 819 TGGGGTTGAACAAATTGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGCTCATCA 878  
|||||  
Db 1018 TGGATTTAACGATGTTTATCTTGGCTTACCAACACTTATCAATCAAAATGGTGCAGTTAA 959  
  
QY 879 AGTGGCTGAACCTTCACTCGATGAGAAGGAATAGAAATTGATGGAATAAATCAG 931  
|||||  
Db 958 AATTATGAACACCACTTAATGATAACGAATACTAATTAAGTAAATCAG 906

RESULT 11  
AF269753  
LOCUS  
DEFINITION  
Staphylococcus epidermidis strain SR1 clone step.1022a07 genomic  
sequence.  
ACCESSION  
AF269753  
VERSION  
AF269753.1 GI:9623650  
KEYWORDS  
SOURCE  
ORGANISM  
Staphylococcus epidermidis.  
Staphylococcus epidermidis  
Bacteria; Firmicutes; Bacillales; Staphylococcus.  
REFERENCE  
1 (bases 1 to 4093)  
AUTHORS  
Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,  
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,  
Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,  
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and  
Furdon,P.J.  
TITLE  
Transposon-mediated sequencing of the Staphylococcus epidermidis  
genome  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 4093)  
AUTHORS  
Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,  
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,  
Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,  
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and  
Furdon,P.J.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (22-MAY-2000) Departments of Genomic Sciences and  
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore  
Drive, Research Triangle Park, North Carolina 27709-3398, USA  
FEATURES  
Location/Qualifiers  
source  
1..4093  
/organism="Staphylococcus epidermidis"  
/strain="SR1"  
/db xref="taxon:1282"  
/clone="step.1022a07"  
BASE COUNT  
1457 a 565 c 765 g 1306 t  
ORIGIN  
Query Match  
18.9%; Score 183.8; DB 1; Length 4093;

Best Local Similarity 53.0%; Pred. No. 2.4e-42;  
Matches 473; Conservative 0; Mismatches 402; Indels 18; Gaps 3;  
  
QY 51 AAAAGTTGTGTAGTGGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCT 110  
|||||  
Db 1104 AAAAGTTGTGTAGTAGGAGACGGTCCGTAGGTTCAAGTTATGCAATTTGCTATGGTGAC 1163  
  
QY 111 CAGCGGCATCGTTTCCGAGATTGTCTCTTATCGACGTGAACAAAGACAGAGGGTGA 170  
|||||  
Db 1164 TCAAGGAATTGCAGATGAATTTGTAATTATTGATATTGCAAAAGATAAAGTGAAGCAGA 1223  
  
QY 171 AAGCATGGAATTAACCCAGCGACACTTCAAATACAAGTCTCG-----AGCGGG 221  
|||||  
Db 1224 CGTTAAAGATTAAACCATGGTGCACTTTACAGTTCTTCCACAGTGAATGTAAGAGCTGG 1283  
  
QY 222 TGATTATCTGACTGCGCTGGCGGGCCCATTTGTTATTGTACATGTGGGATTAAACCAAAA 281  
|||||  
Db 1284 AGAATATGAAGATTGTAAGATGAGATTTAGTTGTTATTACAGAGGTGCACCTCAAAA 1343  
  
QY 282 AAATGGACAAAACAAGGATGGATGCTGTGTCGCAAAAAATGCCAATATGCTGGAATCAT 341  
|||||  
Db 1344 ACCGGGTGAAACTCGTTTACAACCTGTTGAGAAAAATACTAAAAATCATGAAAGTATCGT 1403  
  
QY 342 CCCCAATGTTGCCAAATATGCTCCTGTATACCATCCTGCTTATGCCACCAATCCTGTGCGA 401  
|||||  
Db 1404 AACTAGTGTATGGATAGTGGCTTTGATGGTTCTTCTTAATTTGCTGCAAAACCCAGTTGA 1463  
  
QY 402 TGTTTTGACCTATATTAGTATAAGGCGCTCAGGGTTTCCACTAAGCAGAGTTATCGGCTC 461  
|||||  
Db 1464 TATCTTAACACAGTTATGTTAAAGAAGTTACAGGTTTACCAGCTGAACGTTGTTATTGGTTC 1523  
  
QY 462 AGGTACAGTTCTGGATACTGCTCGTGTGTTTAAATACATCCTCGGAGAGCACTTCAAGATCTC 521  
|||||  
Db 1524 TGGTACAGTGTGATAGTGAAGATTTCAGATATTTAATAAGTAAAGAAATTAGGTGTTAC 1583  
  
QY 522 ATCGGACAGCATCGATGCTGCTGTAATTGGAGAACATGTTGATTCGGGTGTCCTGTCTG 581  
|||||  
Db 1584 ATCAAGTAGTGTTCACGCTAGCATTATAGGTGAACATGGTGAATCTGTAACCTTGCAGTTTG 1643  
  
QY 582 GTCTCTTACCAACATCGACGGGATGAAGCTCCGGGATTACTGCGAAAAAGCCAAACCACAT 641  
|||||  
Db 1644 GTCTCAAGCAACGTTGGAGGTATTTCAAGTATGATACATTTGAAAGAAGAAACTGGTAG 1703  
  
QY 642 ATTTGATCAGAAATGCGTTCCATAGAAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATAT 701  
|||||  
Db 1704 CGATGCTAA-----AGCGAATGAAATTTATATTAATACAAGAGATGCTGCTTACGATAT 1757  
  
QY 702 CATCAAGCGCAAGGCTATACTTCATATGGAATCGACGGGATTACTTCGCATAGTAAA 761  
|||||  
Db 1758 CATTCAGCTAAAGGATCTACGTATTATGGTATAGCTCTAGCACTATTACGTATTTCTAA 1817  
  
QY 762 GCGGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAA---CCGTTGGTGATTATTT 818  
|||||  
Db 1818 AGCTTTACTAAATAATGAAATAGTATTTTGACAGTTTCTAGTCAACTTAATGGTCAATA 1877  
  
QY 819 TGGGGTTGAACAAATTGCTATAAGCGTCCCTACCAAACTCAATAAAAAAGTGGGCTCATCA 878  
|||||  
Db 1878 TGGATTTAACGATGTTTATCTTGGCTTACCAACACTTATCAATCAAAATGGTGCAGTTAA 1937  
  
QY 879 AGTGGCTGAACCTTTCACCTCGATGAGAAGGAATAGAAATTGATGGAATAAATCAG 931  
|||||  
Db 1938 AATTATGAACACCACTTAATTAATGATAACGAATACTAATTAAGTAAATAAATCAG 1990

RESULT 12  
AX145071  
LOCUS  
DEFINITION  
Sequence 3793 from Patent WO0134809.  
ACCESSION  
AX145071  
VERSION  
AX145071.1 GI:14283636  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
synthetic construct





QY 645 TGATCAGAAATCGCTTCCATAGAATCTTTTGAGCAAAACGGAGACGCTGCTTACGATATCAT 704  
Db 919 TAGAAAAGAGGATTAGACAATATCTTTGTTAATGTTCTGATGCGGCATATCAAAATCAT 978  
QY 705 CAAGCGCAAAGCTATATTTCATATGGAATCGCAGCGGATTACTTCGCATAGTAAAGGC 764  
Db 979 TGAGAAAAAGGGCAACGTATTACGGCAATGCAATGGATTAGTCCGTATCACTCGTGC 1038  
QY 765 GATTTTAGAGGATACAGGATCCACACTTACAGTTTCAAC---CGTTGGTGATTATTTGG 821  
Db 1039 TATTTTGCAATGAAAATGCCATCTTAACCGTTTCTGCTCATTTGGACGGCAATATGG 1098  
QY 822 GGTGAACAAATGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGCTCATCAAGT 881  
Db 1099 CGAACGAAATGTTATATTGGCGTGCCTGCCATTATCAACCGAAACGGTATTCGTGAAGT 1158  
QY 882 GGCTGAACCTTTCACCTCGATGAGAAGGAAATAGAA 915  
Db 1159 GATGGAATTGACGCTAAATGAAACAGAACAAACAA 1192

RESULT 14  
TMLDH  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
-JOURNAL  
MEDLINE  
PUBMED  
FEATURES  
source  
Location/Qualifiers  
1. .1130  
/organism="Thermotoga maritima"  
/strain="MSB8 (DSM3109)"  
/db\_xref="taxon:2336"  
/clone="pTLDH25"  
42. .47  
/note="putative"  
65. .70  
/note="putative"  
103. .111  
/note="putative"  
116. .1075  
/gene="ldh"  
116. .1075  
/gene="ldh"  
/EC\_number="1.1.1.27"  
/note="alpha 4 subunit"  
/codon\_start=1  
/transl\_table=11  
/product="L-lactate dehydrogenase"  
/protein\_id="CAA52355.1"  
/db\_xref="GI:396281"  
/db\_xref="SWISS-PROT:P16115"  
/translation="MKIGIVGLGRVGSSTAFALLMKGFAREMVLIDVKKRAEGDALD  
LIHGTFPTRRANIYAGDYADLKGSVDVIVAAGVPQKPGETRLQLLGRNARMKEIARN  
VSKYAPDSIVIVTNPVDVLTYPFLKESGMDPRKVFSGTFLDTLARLRTLIAQHCDFS  
PRSVHVVYIGEHDSEVPVWSGAMIGGIPLQNMCIQKCDSKILENFAEKTGRAAYE

IIERKGATHYAIALAVADIVESIFFDEKRVLTLSVLELYLGVKDLCLISVPVTLKGHG  
VERILELNLNEEELEAFPRKSASILKNAINETABENKHQNTSG"  
BASE COUNT 340 a 238 c 291 g 261 t  
ORIGIN  
Query Match 18.2%; Score 176.6; DB 1; Length 1130;  
Best Local Similarity 52.3%; Pred. No. 2.7e-40;  
Matches 467; Conservative 0; Mismatches 414; Indels 12; Gaps 3;  
QY 50 TAAAGTTGTGGTAGTGGGAGTCGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTC 109  
Db 117 TGAATAATAGGTATCGTAGGACTCGGAAGGGTTGGTCCAGCAGCGCTTTTGCACTCCTGA 176  
QY 110 TCAGCGGCATCGTTCCGAGATTGTCTTATCGACGTGAACAAAGACAAAGCAGAGGGTG 169  
Db 177 TGAAGGTTTTCGAGGGAAATGGTCTGTATAGACGTCTGATAAGAAAAAGAGCAGAAGGAG 236  
QY 170 AAAGCATGGACTTAAACCCAGCAGCACCTTCAAATACAAGGTC-----TCGAGCGGGTG 223  
Db 237 ACGCTCTCGATCTCATTCACGGAACACCTTTCACGAGAGAGCGGAACATCTACGCTGGAG 296  
QY 224 ATTATCCTGACTCGCTGGCGGCCCATTTGTTATTGTACATGTGGGATTAAACCAAAAAA 283  
Db 297 ACTATCGGATCTGAAAGGATCAGATGTGGTGATCGTCGCTCGGGAGTACCTCAGAAAC 356  
QY 284 ATGGACAAACAAGGATGGATCTTCTGTCAAAAAATGCCAACATTATGCTGGAAATCATCC 343  
Db 357 CAGGAGAGACGAGGCTTCAGCTTCTTGGGAGAAATGCAAGAGTGTGAAAGAAATAGCGC 416  
QY 344 CCAATGTTGCCAAATATGCTCTGATACCATCCTGCTTATTGCCACGAATCCTGTGATG 403  
Db 417 GAAACGTCTCCAAATACGCTCCTGATTTCGATTGTTCATCGTGTGTCACGAATCCCCTGATG 476  
QY 404 TTTTGACCTATATTAGTATAAGCGCTCAGGGTTTCCACTAAGACAGAGTTATCGGCTCAG 463  
Db 477 TCCTCACGTATTTCTTCTCAAAAGAGTCCGGAATGGATCCAGGAAGGTGTTTCGGTTCG 536  
QY 464 GTACAGTTCTGGATACCTGCTCGTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCAT 523  
Db 537 GAACGTTCTCGACACAGCAAGGCTCAGAACTTTGTATAGCGCAACACTCGGGCTTTTCTC 596  
QY 524 CGGACAGCATCGATCCCTGTGTAATTGGAGAACATGGTGATTCGGGTGTGCTGTCTGCT 583  
Db 597 CAAGAAGTGTGCAGTGTACGTATAGGAGAACATGGAGATTTCAGAGGTTCTCTGTGTGA 656  
QY 584 CTCTTACCAACATCGACCGGCAATGAAGCTCCGGGATTACTGCGAAAAAGCCACCATAT 643  
Db 657 GCGGTGCCATGATAGGAGGTATCCCCCTACAGAACATGTGTGAGA---TTTGTGAGAAGT 713  
QY 644 TTGATCAGAAATCGTTTCCATAGAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATATCA 703  
Db 714 GCGATTGGAAGATACTGGAAAACTTCGCTGAAAAAACCCAAAAAGAGCCGCGTACGAGATCA 773  
QY 704 TCAAGCGCAAAGGCTATATCTTCATATGGAATCGCAGCGGGATTACTTCGCATAGTAAAGG 763  
Db 774 TAGAGAGAAAGGGAGGCCACCCATTACGCCATAGCTCTCGCCGTGGCCGACATAGTGGAGA 833  
QY 764 CGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAACCGTT---GGTGATTATTTTG 820  
Db 834 GCATCTTCTTCGACGAAAAGAGAGTGTCTTACGCTTTCGGTTTATCTCGAAGATTACCTCG 893  
QY 821 GGGTTGAACAAATTGCTATAAGCGTCCCTACCAAACCTCAATAAAAGTGGGCTCATCAAG 880  
Db 894 GTGTAAAAGATCTCTGTATCAGTGTTCCTGTTACCTTTCGAAAAGCAGCGAGTGGAAGAA 953  
QY 881 TGGCTGAACCTTTCATCTCGATGAGAAGGAAATAGAATTGATGGAAAAATCAGCT 933  
Db 954 TCCTCGAACTGAACCTGAACGAAAGAAAGAACTGGAAGCCTTCAGAAAAATCAGCT 1006

RESULT 15  
AE001823  
LOCUS

AE001823 12762 bp DNA linear BCT 02-JUN-1999



DEFINITION Thermotoga maritima section 135 of 136 of the complete genome.  
ACCESSION AE001823 AE000512  
VERSION AE001823.1 GI:4982441  
KEYWORDS  
SOURCE Thermotoga maritima.  
ORGANISM Thermotoga maritima  
Bacteria; Thermotogae; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.  
REFERENCE 1 (bases 1 to 12762)  
AUTHORS Nelson,K.E., Clayton,R.A., Gill,S.R., Gwinn,M.L., Dodson,R.J., Haft,D.H., Hickey,E.K., Peterson,J.D., Nelson,W.C., Ketchum,K.A., McDonald,L., Utterback,T.R., Malek,J.A., Linher,K.D., Garrett,M.M., Stewart,A.M., Cotton,M.D., Pratt,M.S., Phillips,C.A., Richardson,D., Heidelberg,J., Sutton,G.G., Fleischmann,R.D., Eisen,J.A., Fraser,C.M. et al.  
TITLE Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of Thermotoga maritima  
JOURNAL Nature 399 (6734), 323-329 (1999)  
MEDLINE 99287316  
PUBMED 10360571  
REFERENCE 2 (bases 1 to 12762)  
AUTHORS Nelson,K.E., Clayton,R.A., Gill,S.R., Gwinn,M.L., Dodson,R.J., Haft,D.H., Hickey,E.K., Peterson,J.D., Nelson,W.C., Ketchum,K.A., McDonald,L., Utterback,T.R., Malek,J.A., Linher,K.D., Garrett,M.M., Stewart,A.M., Cotton,M.D., Pratt,M.S., Phillips,C.A., Richardson,D., Heidelberg,J., Sutton,G.G., Fleischmann,R.D., White,O., Salzberg,S.L., Smith,H.O., Venter,J.C. and Fraser,C.M. Direct Submission  
TITLE Submitted (01-JUN-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA  
JOURNAL  
FEATURES  
source 1..12762  
/organism="Thermotoga maritima"  
/db\_xref="taxon:2336"  
complement(134..1660)  
/gene="TM1857"  
complement(134..1660)  
/gene="TM1857"  
/note="similar to GB:AL009126 percent identity: 78.00; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAD36919.1"  
/db\_xref="GI:4982442"  
/translation="MMLWYIVAGAGGLIGYLIANYQINOKLRKAKEDAQTIIEKAEK EANEIKKAIIEGREVEHRLREFEFEKERSRREELRALEREELLRKEELLTRKEENLEK REQVEELKANLEEKMRVEVEKEKRIDEELKRLAGMTVEEAARELILEEARQRYEHDLA KLYKEMKEQVEEEAEKAVIAFAVQRYADPYGEITVSTVSLPDDMKGRIIGREG RNIRTFEKITGVDLIIDDTPEVVVLSFNPRLREIARITLEKLVADGRIHPARIEEMY EKAKQVEKAIKEAGQEATFPKAGVMGLHPVLKLLGKLKYRTSYGQNVLNHSIEVALL AGYMASSELGLNADKARRGGLLHDIGKAVDQELGSHTTIGAEIARRYGEKEDIINMIL SHHGEPEPTPEAVLVAAADALSAAAPGARRESLENYIKRLMKLEIEIAKSFKYVEKAY AIQAGREIRVIVEPDKVDALAEKLAYDISKIBEELEYPGVLKVVVIREKRSVAYAK"  
complement(1657..2142)  
/gene="TM1858"  
complement(1657..2142)  
/gene="TM1858"  
/note="similar to PID:1743349 percent identity: 50.00; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="recX protein, putative"  
/protein\_id="AAD36920.1"  
/db\_xref="GI:4982443"  
/translation="MDYYQWRKKNRCKKRNQLQAKKPLNYALRLKLYRVFEDELRERL KQGFADEEVESTITTLKQGYLDDKAAYLFALDEMRLKLFPRVVKMKLSLGVDE EIIERAIEKALEEIDFHEELKRLKGRFKDRWELRDYLYRRGFDSSLIEEILNKIDGGE E"  
complement(2072..3142)  
/gene="TM1859"  
complement(2072..3142)

/gene="TM1859"  
/note="similar to GB:L23425 SP:P36203 PID:385170  
GB:AE000512 percent identity: 100.00; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="DNA repair protein"  
/protein\_id="AAD36921.1"  
/db\_xref="GI:4982444"  
/translation="MPEEKQKSVLEKALKRIEENFGKGSIMILGDETQVQVVEVIPT GSLAIDIATGVGGRGRIVEIFQEESGKTTALHAIAEAQMGVAAAFIDAEHALD PVYAKNLGVDLKSLLISQPDHGEQALEIVDELVRSGVVDLIVVDSVAALVPAEIEGA MGDMOVGLQARLMSQALRKIAGSVNKSQAVVFTNQIRMKIGVMFGSPETTTGGIALK FYATMRMEVRRGEPIKEGKDVIGNVISVKIVKNKVAAPPFKTAQTYIYIGKIDREYEL FNIATNEGIVDRKGSWYYTTLKGEEVSLGQSSNAVQFLKDNPEIAGEIERIREKY GLLSVEKEEQRKEKSSGEAS"  
complement(3135..3698)  
/gene="TM1860"  
complement(3135..3698)  
/gene="TM1860"  
/note="similar to GB:Pyro\_h percent identity: 54.86; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAD36922.1"  
/db\_xref="GI:4982445"  
/translation="MRTFIAIDVNEEVKKQASEIIEKLMKRGFGATWVSENNHULTLF FLGEVDEQKISETIAEHLCCRVRGFPSPSFVTKGFGYFKRKMSPRVFWLGVENTDRIMK LYEELRNELSHGFSFEKFPVPHITIGRVKYYPDKWEKLIEDIDFPPIEVAVDRPKIY SSTLTPTGPIYKVLVECFEGGLIRYA"  
complement(3702..4238)  
/gene="TM1861"  
complement(3702..4238)  
/gene="TM1861"  
/note="similar to SP:P46322 PID:893358 PID:1842439  
GB:AL009126 percent identity: 60.87; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase"  
/protein\_id="AAD36923.1"  
/db\_xref="GI:4982446"  
/translation="MGVSDMNLANFFSLRLAALVIPVWVFWMEGWISLSFLIFVFAAF TDYLDGFFARKKNQVDFGKVFQVSDKILVITAVAMDVLPLWYVLVVFARDTFVN GLRILAAASRGNVVPARWIGKAKTVSQFVVLIAAYLFKMGFSLNALLMFFVVLSTVTV ISWITYTIDIARITKLEG"  
complement(4220..5512)  
/gene="TM1862"  
complement(4220..5512)  
/gene="TM1862"  
/note="similar to GB:AE000657 percent identity: 67.63; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAD36924.1"  
/db\_xref="GI:4982447"  
/translation="MRVGIKVLGCPKNEADCEVLAVLREGGHEIVFDVKDADVVLVD TCAFIEDAKRESIDEIFSVDKADQYGYKLVKGLVQRYEELKKEIPEVDQWIGVA DPEEIANAIENGTDLPDQPETVYRYRKRIDLBERPYAYVKISDGCDRGCTFCSPSF KGSLSRSIEDITREVEDLLKEGKEIILVAQDTSYGIDLYRKQALPDLLRLNSLN GEFWIRVMYLHPDHLTEEIIISAMLELDKVVPVQHGSDKILKLMGRTRKSSSELK KMLSSIRERFPDAVLRTSIIIVGPPGETEEDFEELKQFVEEIQFDKLGAFYVSDDEGTV AFNLKEKVDPEMAKRQREELLQLQAEISNSRLDRFVGKKLFLVEGKEGKFLVGRTWT EAPEVDGVVFRGKGIQDFLEVIKEHDEYDMWGSVI"  
complement(5509..6039)  
/gene="TM1863"  
complement(5509..6039)  
/gene="TM1863"  
/note="similar to GB:AE000657 percent identity: 66.67; identified by sequence similarity; putative"

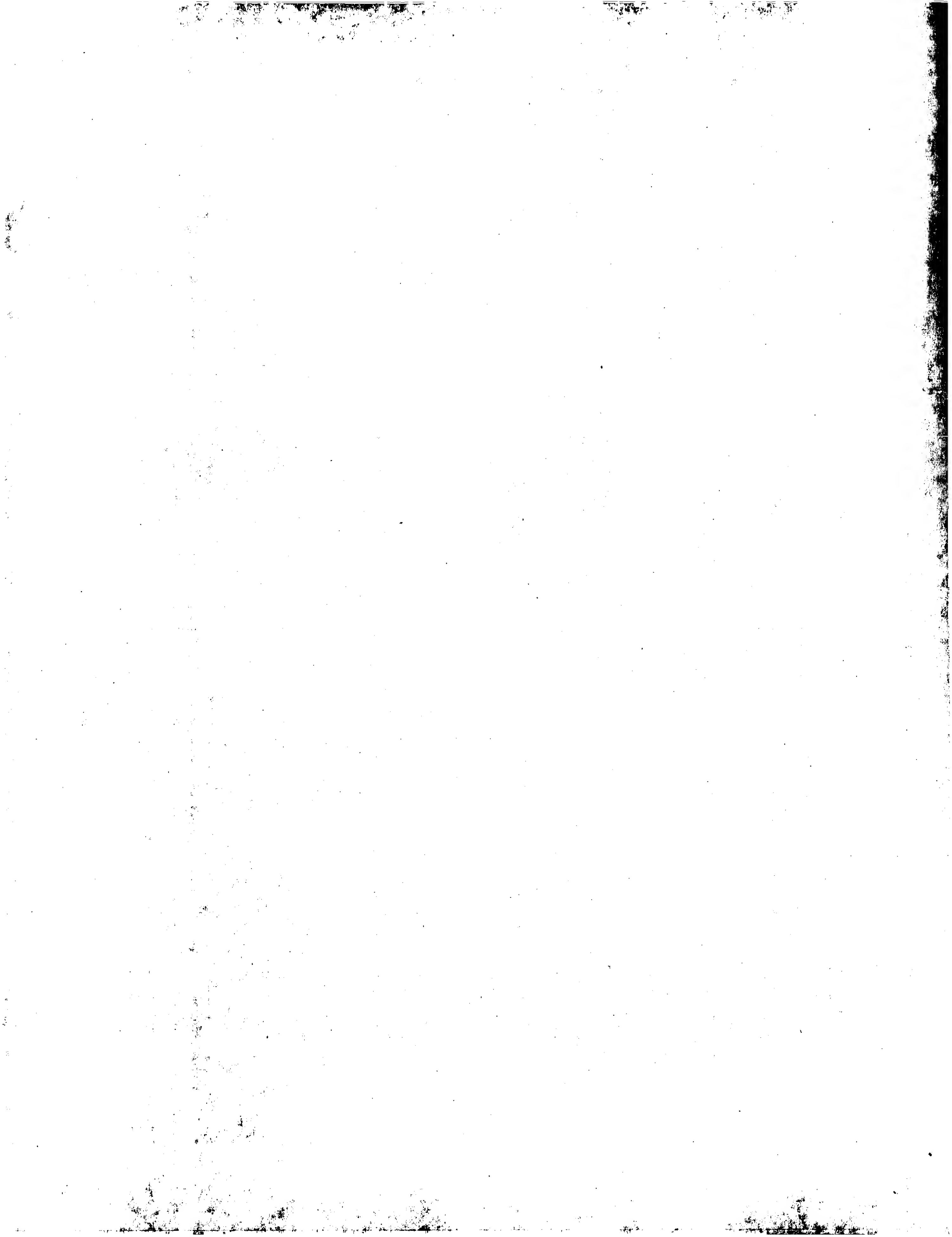
```

/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAD36925.1"
/db_xref="GI:4982448"
/translation="MKVPDMVNLVMPFIFASHIDYWFNEVRPVLGRFGPMDYISDVLD
FEKXLYLYSEMGQGLKGRLVSEFRLVHPFQLADIKRETNEIEKMFSEIGKRKNMNDP
GYIHTQFVLASTKHGWNRIYIGKIYAEVTLVYVRGEFRHMEFTYPNYREEEYKKHL
EKIRELYLKKRRKMK"
complement(6036..6611)
/gene="TM1864"
complement(6036..6611)
/gene="TM1864"
/note="similar to percent identity: 0.00; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAD36926.1"
/db_xref="GI:4982449"
/translation="MSEKWKELGETFRKKREERRITLLDASLFTNINPSKLKRIEEGD
LKGLDAEVIKSYIKRYSFLELSPDEMLKLYEEGKEBEVAEEVEKKPKRKKKEKTR
DLVMFFFLIAGLVFLFSAVENLKLROTPPAYLVAPETIVNGKSVSGEIPLOGEYI
VESRSDVVLKTASEEWTVKIRKFEVSVSEK"
complement(6683..7360)
/gene="TM1865"
complement(6683..7360)
/gene="TM1865"
/note="similar to SP:P32679 PID:409792 GB:U00096
PID:1790431 percent identity: 65.92; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="endonuclease V"
/protein_id="AAD36927.1"
/db_xref="GI:4982450"
/translation="MDYRQLHRWDLPPAEAIVQNELRKKIKLTPYEGEPEVVAGVDL
SPGKEGLAVIVVLEYPFKILEVVSERGEITFPYIPGLLAFREGPLFLKAWEKLR
KPDVVVFDGQGLAHPRLKGIASHMGLFIEIPTIGVAKSRLYGTFKMPEDKRCWSYLY
DGEIIGCVIRTKESGAPIFVSPGHLMDVLESSKRLIKAPTLPGRRIPETRLAHITYQ
RLKKGLF"
complement(7365..8738)
/gene="TM1866"
complement(7365..8738)
/gene="TM1866"
/note="similar to GB:M87645 SP:Q02115 PID:143156
PID:405621 GB:AL009126 percent identity: 52.26; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="membrane bound protein LytR, putative"
/protein_id="AAD36928.1"
/db_xref="GI:4982451"
```

Query Match		18.2%;	Score 176.6;	DB 1;	Length 12762;
Best Local Similarity		52.3%;	Pred. No. 3.5e-40;		
Matches	467;	Conservative	0;	Mismatches	414;
				Indels	12;
				Gaps	3;
Qy	50	TAAAAGTTGTGGTAGTGGGAGTGGGAAGTGTGGTCTGCCACAGCGGTATACGTTGCTTC	109		
Db	8603	TGAAATAGGTATCGTAGGACTCGGAAGGGTGGTTCAGCACGGCTTTTGCACTCCTGA	8662		
Qy	110	TCAGCGGCATCGTTTCCGAGATTGTCTTATCGACGTGAACAAAGACAAAGCAGAGGGTG	169		
Db	8663	TGAAAGTTTTCGAGGGAAATGGTCTGATAGACGTGCGATAAGAAAGACAGAGGAG	8722		
Qy	170	AAAGCATGGACTTAAACCACGACGACACCTTCAAATACAAAGTC-----TCGAGCGGGTG	223		
Db	8723	ACGCTCTCGATCTCATTCACGGAACACCTTTCACGAGAGCGGAACATCTACGCTGGAG	8782		
Qy	224	ATTATCCTGACTGCGCTGGCGGCGCCATTGTTATTGTGCATGTGGGATTAAACCAAAAAA	283		
Db	8783	ACTATCGGATCTGAAAGGATCAGATGTGGTGATCGTCTGCGGGAGTACCTCAGAAAC	8842		

Qy	284	ATGACAAAACAAGGATGGATCTTGCTGCAAAAAATGCCAACATTATGCTGGAAATCATCC	343
Db	8843	CAGAGAGACGAGGCTTCAGCTTCTTGGGAGAAATGCAAGAGTGATGAAAGAAATAGCGC	8902
Qy	344	CCAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTCGCCACGAATCCTGTGATG	403
Db	8903	GAAACGTCTCCAAATACGCTCCTGATTGATTGTCATCGTGGTCAGAAATCCCCTCGATG	8962
Qy	404	TTTTGACCTATATTAGCTATAAGCGCTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCAG	463
Db	8963	TCCTCACGTATTCTTCTCAAGAGTCCGGAATGGATCCCAGGAAGGTGTTTCGGTTCCG	9022
Qy	464	GTACAGTTCTGGATACTGCTCGTTTAAATACATCCTCGGAGAGCACCTTCAAGATCTCAT	523
Db	9023	GAAACGGTCTCGACACAGCAAGGCTCAGAACTTTGATAGCGCAACACACTCGGGCTTTTCTC	9082
Qy	524	CGGACAGCATCGATGCCCTGTGTAATTGGAGAACATGGTGATTCCGGTGTGCTCTGGT	583
Db	9083	CAAGAAGTGTGCACGTGTACGTGATAGGAGAACATGGAGATTTCAGAGGTTCTGTGTGGA	9142
Qy	584	CTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTCGGAAAAAGCCAACCATAT	643
Db	9143	GCGGTGCCATGATAGGAGGTATCCCCCTACAGAACATGTGTGAGA---TTTGTCAAGAGT	9199
Qy	644	TTGATCAGAAATGCGTTCCATAGAAATCTTTGAGCAACCGCGAGACGCTGCTTACGATATCA	703
Db	9200	GCGATTGGAAGATACTGGAAAACTTCGCTGAAAAAACCAAAAGAGCCGCGTACGAGATCA	9259
Qy	704	TCAAGCGCAAAGGCTATATTTCATATGGAATCGCAGCGGGATTACTTCGCATAGTAAAGG	763
Db	9260	TAGAGAGAAAGGGAGCCACCATTACGCCATAGCTCTCGCCGTGGCCGACATAGTGGAGA	9319
Qy	764	CGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAACCGTT--GGTGATTATTTTG	820
Db	9320	GCATCTTCTTCGACGAAAAAGAGAGTGCTTACGCTTTCGCTTTTATCTCGAAGATTACCTCG	9379
Qy	821	GGTTGAACAAATTGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGGCTCATCAAG	880
Db	9380	GTGTAAGAGATCTCTGTATCAGTGTTCTCTGTTTACCCCTTGGAAAGCACGGAGTGGAAAGAA	9439
Qy	881	TGGCTGAACCTTTCACTCGATGAGAGGAATAAGATTGATGGAAAAATCAGCT	933
Db	9440	TCCTCGAACTGAACCTGAACGAAAGAACTGGAGCCTTCAGAAAGTCAGCT	9492

Search completed: July 31, 2003, 09:42:47  
Job time : 2773 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 06:50:56 ; Search time 278 Seconds  
(without alignments)  
7873.893 Million cell updates/sec

Title: US-09-992-430B-21  
Perfect score: 972  
Sequence: 1 atgttccaagatacaaaagtc.....ttgagcatctggagatcaat 972

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*

2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*

3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*

4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*

5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*

6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*

7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*

8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*

9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*

10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*

11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*

12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*

13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*

14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*

15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*

16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*

17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*

18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*

19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*

20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*

21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*

22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*

23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*

24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	205.2	21.1	2365589	24 ABA90521	Genomic sequence o
2	186.2	19.2	28690	20 AAX13075	Enterococcus faeca
3	183.8	18.9	969	22 AAH53009	S. epidermidis ope
4	183.8	18.9	969	22 AAH53495	S. epidermidis ope
5	183.8	18.9	993	24 ABA90521	Staphylococcus epi
C 6	183.8	18.9	3019	22 AAH54991	S. epidermidis gen
7	183.8	18.9	4093	22 AAH54429	S. epidermidis gen
8	183.6	18.9	1255	24 AAD37417	Bacillus strain LN
C 9	173	17.8	13154	20 AAX13275	Enterococcus faeca

10	172.6	17.8	3264	18 AAV74735	Staphylococcus aur
11	157.2	16.2	981	24 ABN68374	Streptococcus poly
12	155.8	16.0	945	20 AAZ09139	B. flavum lactate
13	155.8	16.0	1065	22 AAF71424	Corynebacterium gl
14	154.2	15.9	942	22 AAH68175	C glutamicum codin
C 15	154.2	15.9	309400	22 AAH68534	C glutamicum codin
C 16	154.2	15.9	1163020	24 ABQ67197	Listeria innocua c
C 17	154.2	15.9	3011208	24 ABQ69245	Listeria innocua D
C 18	153.6	15.8	5278	19 AAV52360	Streptococcus pneu
C 19	153.4	15.8	2944528	24 ABA03041	Listeria monocytog
20	151	15.5	747	24 ABK72782	Bacillus lichenifo
21	149	15.3	1197	12 AAQ10162	L-lactic acid dehy
C 22	147.2	15.1	5449	21 AAC64598	Streptococcus pneu
C 23	146.2	15.0	14280	18 AAV74368	Staphylococcus aur
24	142.6	14.7	3073	12 AAQ15280	L-lactic acid dehy
25	138.8	14.3	1234	9 AAN80494	Sequence encoding
26	136	14.0	1680	24 AAS17494	Human cDNA encodin
27	136	14.0	1759	22 AAH64738	Human secreted pro
28	133.4	13.7	910715	20 AAX20248	Borrelia burgdorfe
29	132.8	13.7	987	24 ABN68373	Streptococcus poly
C 30	132.8	13.7	2155561	24 ABN71527	Streptococcus poly
C 31	129.2	13.3	2621	24 ABQ70916	Listeria monocytog
32	128.4	13.2	1656	20 AAX97742	Extended human sec
33	125	12.9	1661	21 AAA09330	Human cancer assoc
34	125	12.9	1661	24 ABK84497	Human cDNA differe
35	125	12.9	1732	22 AAI59428	Human polynucleoti
36	125	12.9	1792	22 AAI61214	Human polynucleoti
37	125	12.9	2223	22 AAH72820	Human cervical can
38	125	12.9	2224	22 AAH72672	Human cervical can
39	123.6	12.7	1261	21 AAC42568	Arabidopsis thalia
40	121.8	12.5	950	20 AAX61751	B. burgdorferi ant
41	121.2	12.5	4913	24 AAD37416	Bacillus strain TN
42	120.8	12.4	1755	22 AAH64739	Human secreted pro
43	118.8	12.2	1519	24 ABQ69016	Listeria monocytog
44	116.6	12.0	29559	23 AAS59546	Propionibacterium
45	115.8	11.9	887	20 AAX61752	B. burgdorferi ant

ALIGNMENTS

RESULT 1

ABA90521/c

ID ABA90521 standard; DNA; 2365589 BP.

XX ABA90521;

AC ABA90521;

XX ABA90521;

DT 16-MAY-2002 (first entry)

XX Genomic sequence of Lactococcus lactis IL1403.

DE Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese; ds.

XX Lactococcus lactis IL1403.

OS Lactococcus lactis IL1403.

XX Lactococcus lactis IL1403.

PN FR2807446-A1.

XX 12-OCT-2001.

PD 11-APR-2000; 2000FR-0004630.

XX 11-APR-2000; 2000FR-0004630.

PF 11-APR-2000; 2000FR-0004630.

XX 11-APR-2000; 2000FR-0004630.

XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.

PA Bolotine A, Sorokine A, Renault P, Ehrlich SD;

XX WPI; 2002-043418/06.

XX New nucleotide sequence useful in the identification or Lactococcus

PT lactis and related species -

XX Claim 1; SEQ ID 1; 2504pp; French.



The present invention is related to a *Lactococcus lactis* nucleotide sequence (AB90521) and related proteins (AB53300-AB55621). The nucleic acid sequence is useful in the detection and/or amplification of nucleic acid sequence, particularly to identify *Lactococcus lactis* or related species. The proteins of the invention are useful for the biosynthesis or biodegradation of a composition of interest. The invention helps research in lactic bacteria, particularly useful in the production of yogurt and cheese.  
 Note: The sequence data for this patent is based on equivalent patent WO200177334 (published 18-Oct-2001) which is available in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

Sequence	2365589 BP; 765914 A; 415261 C; 420487 G; 763927 T; 0 other;
----------	--

Query Match	21.1%;	Score	205.2;	DB	24;	Length	2365589;
Best Local Similarity	53.4%;	Pred. No.	1.6e-52;				
Matches	507;	Conservative	0;	Mismatches	428;	Indels	15; Gaps 3;

QY	29	TAAGAACTGATGCCAAACACGATAAAAGTTGTGGTAGTGGGAGTGGGAAGTGTGGGTCTG	88
DB	380606	TAAAAATCAAAATTACAAGCAGAAAAGTAGTTGTTATAGGAACAGGTTTGTGGAAACAA	380547
QY	89	CCACAGCGTATACGTTGCTTCTCAGCGGCATCGTTTCCGAGATTGTCCTTATCGACGTA	148
DB	380546	GTATTGCTTATTCAATGATTAAATCAAGGACTTGTTAATGAGTTGGTTTGATAGATGTTA	380487
QY	149	ACAAAGCAAAAGCAGAGGTTGAAAGCATGGACTTAAACACGACGACCTTC-----	200
DB	380486	ATCAGGATAAAGCTGAGGTTGAAAGCCTTAGATTATTAGATGGCGTTCTTGGGGCAAG	380427
QY	201	-AAATACAAGGTCTCGAGCGGTGATTATCCTGACTGCGCTGGCGGCCCATTTGTTATTG	259
DB	380426	AAAATGTTATTGTGAGAGCTGGAGATTATAAGGATTGTAAAAACGCTGATATTGTGTCG	380367
QY	260	TCACATGTGGGATTAAACCAAAAAATGGACAAACAGGATGGATCTTGCTGCAAAAAATG	319
DB	380366	TTACGGTGGGTAAATCAAAAACAGGTCAATCACGTTTAGATTGGTTAATACAAATG	380307
QY	320	CCAACATTATGCTGGAAATCATCCCAATGTTGGCCAAATATGCTCCTGATACCATCCTGC	379
DB	380306	CTAAAAATTATGCGCTCAATCGTGACTCAGGTTATGGACTCAGGTTTGTGATGTTTGTG	380247
QY	380	TTATTGCCAGAAATCCTGTGATGTTTTTGACCTATATTAGCTATAAGCGCTCAGGTTTC	439
DB	380246	TAATTGCTTCAAAATCCTGTGGATATTCTTACCTATGTCGCTTGGGAAACTTCTGGTCTTG	380187
QY	440	CACTAAGCAGAGTTATCGGCTCAGGTACAGTTCTGGATACATGCTCGTTTAAATACATCC	499
DB	380186	ATCAATCTCGAATTGTGGAATCTGGAACCTGTACAACCTTGGATACCTACTCGTTTAGAAAAGAAAT	380127
QY	500	TCGGAGAGCACTTCAAGATCTCATCGGACAGCATCGATGCCTGTGTAATTGGAGAACATG	559
DB	380126	TGGCCACAAAGCTAGAAATTGATCCACGGAGCGTTTCATGGATATATTATTGGCGAACATG	380067
QY	560	GTGATTCCGGTGTGCCTGTCTGCTCTTCTTACCAACATCGACGGCATGAAGCTCCGGATT	619
DB	380066	GCGATTCTGAAGTTGCAGTTTGGTCACACACGACGAGTGGTGGAAAACCTATTCTTGAAT	380007
QY	620	ACTGCGAAAAGCCAAACCATATTTTGATCAGAAATCGGTTCCATAGAAATCTTTGAGCAA	679
DB	380006	TTATTGTTAAAAATAAAAAATAGGGTTGAAGATTGTTCTAATTTGTCAAATAAGGTAA	379947
QY	680	CGCGAGACGTGCTTACGATATCATCAAGCGCAAGGCTATACCTCATPATGGAATCGCAG	739
DB	379946	---AAAATGCTGCCTATGAGATTATTGATAAAAAACAGGCTACTTATTACGGCATTTGAA	379890
QY	740	CGGGATTACTTCGCATAGTAAAGGCGATTTTTAGAGGATACAGGATCCACACTTACAGTTT	799
DB	379889	TGAGTACAGCTAGAAATTGTCAAGGCCATCTTAAATAATGAACAAGCAATCTACCAGTTT	379830
QY	800	CAACC---GTTGGTGATTATTTTGGGGTTGAACAAATTGCTATAAGCGTCCCTACCAAC	856

[illegible]

QY 111 CAGCGGCATCGTTTCCGAGATTGTCTTATCGACGTGAACAAAGACAAAGCAGAGGTGA 170  
| | | | |  
Db 16670 CCAAGGATTGCGAATGAATTAATCTTAGTTGATATTGACAAAGCCAAATCTGAAGCGGA 16729  
| | | | |  
QY 171 AAGCATGGACTTAAACCACG-----CAGCACCTTCAAAATACAGGTCTCGAGCGGTT 222  
| | | | |  
Db 16730 NCAATTGACTTATTAGATTGGTGKCYTGGGKCMAGAAATGTAAACGTCTGGGCTGGC 16789  
| | | | |  
QY 223 GATTATCCTGACTGCGCTGGCGCGGCCATTGTTATTGTACATGTGGGATTAACCAAAAA 282  
| | | | |  
Db 16790 GACTATCAAGACTGCCAAGATGCCGATATCGTCGTGATTACAGCTGGCGCTAATCAAAA 16849  
| | | | |  
QY 283 AATGGACAAACAGGATGGATCTTGCTGCAAAAAATGCCAACATATATGCTGGAATCATC 342  
| | | | |  
Db 16850 CCTGGGCAAGTCGTCTAGATTGGTTTCAATCAATGCAGAAATATGAAAACAATTTGTT 16909  
| | | | |  
QY 343 CCCAATGTTGCCAAATATGCTCCTGATACCATCTCTGCTTATTGCCACGAATCCTGTCGAT 402  
| | | | |  
Db 16910 AACAAATATCATGAAATCTGGTTTGTAGGAATTTTAGTGATTGCTCAAAATCCTGTGCGAT 16969  
| | | | |  
QY 403 GTTTTGACCTATATTAGCTATAAGCGCTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCA 462  
| | | | |  
Db 16970 GTACTGACTTATGTGGCCTGGCAAGCTTCTGGTTTACCTGTTTCAAGAGTAATTGGAAT 17029  
| | | | |  
QY 463 GGTACAGTTCTGGATACTGCTCGTTTAAATACATCCTCGGAGGACATTCAAGATCTCA 522  
| | | | |  
Db 17030 GGTACAACTTTGGACACAACCTCGTTTCCGCAAGAACTGTCTCAAGCTTTAGCGATTGAT 17089  
| | | | |  
QY 523 TCGGACAGCATCGATGCCTGTGTAATTGGAGAAACATGGTGATTCGGGTGTGCTGTCTGG 582  
| | | | |  
Db 17090 CCACGCAATGTTACGGCTATATTATTGGCGAAACACGGGGATTCTGAAGTGGCTGTCTGG 17149  
| | | | |  
QY 583 TCTCTTACCAACATCGACGGCATGAAGTCCGGGATTACTGCGAAAAAGCCAAACCACTA 642  
| | | | |  
Db 17150 TCTCATACCATGATTGGTACCAACCTATTTTAGAAATTTGTGGATACGACAGCGCTTA 17209  
| | | | |  
QY 643 TTTGATCAGAATCGTTTCCATAGAACTTTTGAGCAACCGGAGACGCTGCTTACGATATC 702  
| | | | |  
Db 17210 ACTAGTGACGAT--TTACCAATCATTTCTGATAAAGTGAAAAATACAGCTTATGAAAT 17266  
| | | | |  
QY 703 ATCAAGCGCAAGGCTATACTTCAATATGGAATCGCAGCGGGATTACTTCGCATAGTAAAG 762  
| | | | |  
Db 17267 ATCGATCGCAACAAAGCGACCTATTATGGGATTGGTATGAGTACTGCACGCAATTGTTAA 17326  
| | | | |  
QY 763 GCGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAAC---CGTTGGTGATTATTTT 819  
| | | | |  
Db 17327 GCCATTTTAAATAATGAACAAGCTATTTTACCTGTCTCAGCTTACTTAGATGGGCAATAT 17386  
| | | | |  
QY 820 GGGGTTGAACAAATTGCTATAAGCGTCCCTACCAACTCAATAAAGTGGGCTCATCAA 879  
| | | | |  
Db 17387 GGGCAACAAGATGATTTTACAGGGATTCTCTGCAGTCGTTGGCAATCAAGGTGTGACTGAC 17446  
| | | | |  
QY 880 GTGGCTGAACCTTTCACTCGATGAGAAAGGAAATAGAATTGATGGAATAAATCAGCTAGTCAG 939  
| | | | |  
Db 17447 ATTATTGAATTGAATCTGAATGCCGCTGAAAAAAGAACTCTTCCAAAAAATCAGTGACACAA 17506  
| | | | |  
QY 940 ATCAAAATCAGTGATTG 955  
| | | | |  
Db 17507 TTAACAACAGTGATGG 17522  
| | | | |

RESULT 3  
AAH53009  
ID AAH53009 standard; DNA; 969 BP.  
XX  
AC AAH53009;  
XX  
DT 03-SEP-2001 (first entry)  
XX  
DE S. epidermidis open reading frame nucleotide sequence SEQ ID NO:1411.  
XX  
KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;

KW vaccination; endocarditis; ds.  
XX Staphylococcus epidermidis.  
OS  
XX WO200134809-A2.  
PN  
XX 17-MAY-2001.  
PD  
XX 09-NOV-2000; 2000WO-US30782.  
PF  
XX 09-NOV-1999; 99US-0164258.  
PR  
XX (GLAX ) GLAXO GROUP LTD.  
PA  
XX Kimmerly WJ;  
PI  
XX WPI; 2001-316495/33.  
DR P-PSDB; AAG82159.  
XX  
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
PT useful for vaccinating against infections, e.g. endocarditis -  
PT  
XX Claim 8; Page 398-399; 2188pp; English.  
PS  
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.  
CC (I) and (II) can have antibacterial activity and therefore can be used  
CC in vaccination. The nucleic acids (I) may be used to produce the  
CC S. epidermidis polypeptides (II) via the production of vectors  
CC containing them which are used to produce hosts cells which express the  
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
CC used to vaccinate subjects and to raise antibodies against the bacteria.  
CC The polypeptides may also be used to assay for other inhibitors of their  
CC activity and therefore identify compounds that may be used for the  
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to  
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA  
CC polynucleotide sequences from the present invention. AAH55091 to  
CC AAH55098 represent oligonucleotide sequences and primers which are used  
CC in the exemplification of the present invention.  
CC N.B. The present invention specifically claims all the polynucleotide  
CC sequences given in the sequence listing of the present specification,  
CC however the sequence listing only goes up to SEQ ID NO:4454 so even  
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,  
CC no sequences are present for SEQ ID NO:4455 to 4464.  
XX  
SQ Sequence 969 BP; 339 A; 130 C; 194 G; 306 T; 0 other;

Query Match 18.9%; Score 183.8; DB 22; Length 969;  
Best Local Similarity 53.0%; Pred. No. 1.5e-47;  
Matches 473; Conservative 0; Mismatches 402; Indels 18; Gaps 3;  
QY 51 AAAAGTTGTGGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCT 110  
| | | | |  
Db 36 AAAAGTTGTTTAGTAGGAGACGGTTCCGTTAGTTTCAAGTTATGCTATGCTGTTGAC 95  
| | | | |  
QY 111 CAGCGGCATCGTTTCCGAGATTGTCTTATCGACGTGAACAAAGCAGAGGTGA 170  
| | | | |  
Db 96 TCAAGGAATTGCAGATGAATTGTAATTATTGATATTGCAAAAGATAAAGTGGAGCAGA 155  
| | | | |  
QY 171 AAGCATGGACTTAAACCACGACGACCTTCAAAATACAAAGTCTCG-----AGCGGG 221  
| | | | |  
Db 156 CGTTAAAGATTTAAACCACGATGGTGCACTTTACAGTTCTTACCAGTGAAGTGG 215  
| | | | |  
QY 222 TGATTATCCTGACTGCGCTGGCGCGGCCATTGTTATTGTTCACATGTGGGATTAACCAAAA 281  
| | | | |  
Db 216 AGAATATGAAGATTGTAAGATGCAGATTTAGTTGTTATTATACAGCAGGTGCACCTCAAAA 275  
| | | | |  
QY 282 AAATGGACAAACAAAGGATGGATCTTGCTGCAAAAAATGCCAAATATGCTGGAATCAT 341  
| | | | |  
Db 276 ACCGGGTGAAACTCGTTTACAACTTGTGAGAAAAATACTAAAATCATGAAAAGTATCGT 335  
| | | | |  
QY 342 CCCCAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTGCCACGAATCCTGTGCA 401  
| | | | |

Db 336 AACTAGTGTCAATGATAGTGGCTTTGATGGTTTCTTCTTAATTGCTGCAAAACCCAGTTGA 395  
Qy 402 TGTTTTGACCTATATTAGCTATAAGGCGTACAGGTTTCCACTAAGCAGAGTTATCGGCTC 461  
Db 396 TATCTTAACACGTTATGTTAAAGAAGTTACAGGTTTACCAGCTGAACGTGTTATTGGTTC 455  
Qy 462 AGGTACAGTTCGTGACTGCTGCTGTTTAAATACATCCTCGGAGACGACTTCAAGATCTC 521  
Db 456 TGGTACAGTGTCTGATAGTCAAGATTTCAGATATTAAAGTAAGTAAGTAAGTGGTTTAC 515  
Qy 522 ATCGGACAGCATCGATGCTGTGTGTAATTGGAGAACATGGTGATTCGGGTGCTGCTG 581  
Db 516 ATCAAGTAGTGTTCACGCTAGCATTTATAGGTGAACATGGTGACTCTGAACCTTGCAGTTG 575  
Qy 582 GTCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCAAACCAT 641  
Db 576 GTCTCAAGCAACGTTGGAGGTATTTTCAGTGTATGATACATTTGAAGAAGAACTGGTAG 635  
Qy 642 ATTTGATCAGAAATGCGTTCATAGAAATCTTTGAGCAACGCGAGACGCTGCTTACGATAT 701  
Db 636 CGATGCTAA-----AGCGAATGAAATTTATATTAACAAGAGATGCTGCTTACGATAT 689  
Qy 702 CATCAAGCGCAAGGCTATACTTCAATATGGAATCGCAGCGGATTACTTCGCATAGTAA 761  
Db 690 CATTCAGCTAAAGGATCTACGTATTATGGTATAGCTCTAGCACATATTACGTATTTCTAA 749  
Qy 762 GCGGATTTAGAGATACAGGATCCACACTTACAGTTTCAA---CCGTTGGTGATTATTT 818  
Db 750 AGCTTTTACTAAATGAATAATAGTATTTTGACAGTTTCTAGTCAACTAATAGTCAATA 809  
Qy 819 TGGGGTTGAACAAATGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGCTCATCA 878  
Db 810 TGGATTTAAACGATGTTTATCTTGGCTTACCACACTTATCAATCAAAATGGTGCAGTTAA 869  
Qy 879 AGTGGCTGAACCTTTCACTCGATGAGAAGGAATAAGAAATTGATGGAATAAATCAG 931  
Db 870 AATTTATGAACACCATTAATGAATGATAACGAACCTACAATTTACTAGAAAAATCAG 922

RESULT 4

AAH53495  
ID AAH53495 standard; DNA; 969 BP.  
XX AC AAH53495;  
XX AC  
XX DT 03-SEP-2001 (first entry)  
XX DE  
XX DE S. epidermidis open reading frame nucleotide sequence SEQ ID NO:2383.  
XX KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;  
XX KW vaccination; endocarditis; ds.  
XX OS Staphylococcus epidermidis.  
XX PN WO200134809-A2.  
XX PD 17-MAY-2001.  
XX PD  
XX PF 09-NOV-2000; 2000WO-US30782.  
XX PF  
XX PR 09-NOV-1999; 99US-0164258.  
XX PR  
XX PA (GLAXO ) GLAXO GROUP LTD.  
XX PI  
XX PI Kimmerly WJ;  
XX XX  
XX DR WPI; 2001-316495/33.  
XX DR P-PSDB; AAG82645.  
XX XX  
XX PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
XX PT useful for vaccinating against infections, e.g. endocarditis -  
XX XX  
XX PS Claim 8; Page 629-630; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
CC (II), given in AAC81454 to AAG83120, from Staphylococcus epidermidis.  
CC (I) and (II) can have antibacterial activity and therefore can be used  
CC in vaccination. The nucleic acids (I) may be used to produce the  
CC S. epidermidis polypeptides (II) via the production of vectors  
CC containing them which are used to produce hosts cells which express the  
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
CC used to vaccinate subjects and to raise antibodies against the bacteria.  
CC The polypeptides may also be used to assay for other inhibitors of their  
CC activity and therefore identify compounds that may be used for the  
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to  
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA  
CC polynucleotide sequences from the present invention. AAH55091 to  
CC AAH55098 represent oligonucleotide sequences and primers which are used  
CC in the exemplification of the present invention.  
CC N.B. The present invention specifically claims all the polynucleotide  
CC sequences given in the sequence listing of the present specification,  
CC however the sequence listing only goes up to SEQ ID NO:4454 so even  
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,  
CC no sequences are present for SEQ ID NO:4455 to 4464.  
XX  
SQ Sequence 969 BP; 339 A; 130 C; 194 G; 306 T; 0 other;  
  
Query Match 18.9%; Score 183.8; DB 22; Length 969;  
Best Local Similarity 53.0%; Pred. No. 1.5e-47;  
Matches 473; Conservative 0; Mismatches 402; Indels 18; Gaps 3;  
  
Qy 51 AAAAGTTGTGTAGTGGGAGTGGGAAGTGTGGTCTGCCACAGCGTATACGTTGCTTCT 110  
Db 36 AAAAGTTGTTTAGTAGGAGACGGTTCCTAGGTTCAAGTTATGCATTTGCTATGGTGAC 95  
Qy 111 CAGCGGCATCGTTTCCGAGATTTGTAATTATTGATATTGCAAAAGATAAAGTGGAAAGCAGA 170  
Db 96 TCAAGGAATTGCAGATGAATTTGTAATTATTGATATTGCAAAAGATAAAGTGGAAAGCAGA 155  
Qy 171 AAGCATGGACTTAAACCCAGCAGCACCTTCAAATACAAAGTCTCG-----AGCGGG 221  
Db 156 CGTTAAAGATTTAAACCATGGTGCACCTTTACAGTTCTTCAACAGTGAAGTGTAAAGCTGG 215  
Qy 222 TGATTATCCTGACTGCGCTGGCGGGCCATTGTTATTGTACATGTGGGATTAAACCAAAA 281  
Db 216 AGAATATGAAGATTGTAAGATGCAGATTTAGTTGTTATTACAGCAGGTGCACCTCAAAA 275  
Qy 282 AAATGGACAAACAAGGATGGATCTTCTGCAAAAATGCCAAACATTATGCTGGAATCAT 341  
Db 276 ACCGGTGAAACTCGTTTACAACTTGTGAGAAAAATACTAAATCATGAAAAGTATCGT 335  
Qy 342 CCCCAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTGCCACGAATCCTGTGCA 401  
Db 336 AACTAGTGTCAATGATAGTGGCTTTGATGGTTCTTCTTCTTAATTGCTGCAAAACCCAGTTGA 395  
Qy 402 TGTTTTGACCTATATTAGTATAAGGCGTACAGGTTTCCACTAAGCAGAGTTATCGGCTC 461  
Db 396 TATCTTAACACGTTATGTTAAAGAAGTTACAGGTTTACCAGCTGAACGTGTTATTGGTTC 455  
Qy 462 AGGTACAGTTCGTGATACTGCTCGTTTAAATACATCCTCGGAGAGCACTTCAAGATCTC 521  
Db 456 TGGTACAGTGTCTGATAGTGAAGATTTCAGATATTAAAGTAAGTAAGTAAGTGGTTTAC 515  
Qy 522 ATCGGACAGCATCGATGCTGTGTAATTGGAGAACATGGTGATTCGGGTGCTGCTG 581  
Db 516 ATCAAGTAGTGTTCACGCTAGCATTTATAGGTGAACATGGTGACTCTGAACCTTGCAGTTG 575  
Qy 582 GTCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCAAACCAT 641  
Db 576 GTCTCAAGCAACGTTGGAGGTATTTTCAGTGTATGATACATTTGAAGAAGAACTGGTAG 635  
Qy 642 ATTTGATCAGAAATGCGTTCATAGAAATCTTTGAGCAACGCGAGACGCTGCTTACGATAT 701  
Db 636 CGATGCTAA-----AGCGAATGAAATTTATATTAACAAGAGATGCTGCTTACGATAT 689  
Qy 702 CATCAAGCGCAAGGCTATACTTCAATATGGAATCGCAGCGGATTACTTCGCATAGTAA 761



Db 690 CATTCAAGCTAAAGGATCTACGTATTATGGTATAGCTCTAGCACTATTACGTATTCTTAA 749  
QY 762 GCGGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAA---CCGTTGGTGATTATTT 818  
Db 750 AGCTTTTACTAAATAATGAAATAAGTATTGACAGTTTCTAGTCACTTAATGGTCAATA 809  
QY 819 TGGGGTTGAACAAATTTGCTATAAGCGTCCCTACCAACTCAATAAAGTGGGGCTCATCA 878  
Db 810 TGGATTTAACGATGTTTATCTTGGCTTACCAACACTTATCAATCAAAATGGTGCAGTTAA 869  
QY 879 AGTGGCTGAACCTTTCACTCGATGAGAAGGAATAAGATTGATGGAAAAATCAG 931  
Db 870 AATTATGAACACCATTAAATGATACGAACACTACAATTACTAGAAAAATCAG 922

RESULT 5

ABN92028  
ID ABN92028 standard; DNA; 993 BP.

AC ABN92028;

XX 24-JUL-2002 (first entry)

DT Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:1491.

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
KW antibacterial; gene therapy; gene; ds.

XX Staphylococcus epidermidis.

OS US6380370-B1.

XX 30-APR-2002.

XX 13-AUG-1998; 98US-0134001.

XX 14-AUG-1997; 97US-055779P.

PR 08-NOV-1997; 97US-064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

PI WPI; 2002-381255/41.

DR P-PSDB; ABP39483.

XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis  
PT polypeptide, useful for diagnosing and treating bacterial infections -  
PS Disclosure; SEQ ID 1491; 267pp; English.

XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have  
CC antibacterial activity and can be used in gene therapy. The sequences  
CC can also be used in the diagnosis and treatment of bacterial infections,  
CC particularly S. epidermidis infections. The sequences can be used to  
CC screen for compounds able to interfere with the S. epidermidis life  
CC cycle or inhibit S. epidermidis infection.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC USPTO web site.

XX Sequence 993 BP; 345 A; 133 C; 197 G; 318 T; 0 other;

SQ Query Match 18.9%; Score 183.8; DB 24; Length 993;  
Best Local Similarity 53.0%; Pred. No. 1.5e-47;  
Matches 473; Conservative 0; Mismatches 402; Indels 18; Gaps 3;

QY 51 AAAAGTTGTGGTAGTGGGAGTGGGAAGTGTGGGTCTGCCACACCGGTATACGTTGCTTCT 110  
Db 60 AAAAGTTGTTTAGTAGGAGACGGTTCCTAGGTTCAGTTATGCATTGCTATGGTGAC 119

QY 111 CAGCGGCATCGTTTCCGAGATTGTCTTATCGACGTGAACAAAAGACAAAGCAGAGGGTGA 170  
Db 120 TCAAGGAATTCAGATGAATTTGTAATTATTGATATTGCAAAAAGATAAAGTGAAGCAGA 179  
QY 171 AAGCATGGACTTAAACCCAGCAGCACCTTCAAATACAAAGGTCTCG-----AGCGGG 221  
Db 180 CGTTAAAGATTTAAACCATGGTGCACCTTTACAGTTCTTACCAGTGAAGTAAAGCTGG 239  
QY 222 TGATTATCCTGACTGGCTGGCGCGGCCCATTTGTTATTGTACATGTGGGATTAAACCAAAA 281  
Db 240 AGAATATGAAGATTGTAAGATGCAGATTTAGTTGTTATTACAGCAGGTGCACCTCAAAA 299  
QY 282 AAATGGACAAAACAGGATGGATCTTTGCTGCAAAAAATGCCAACATTATGCTGGAATCAT 341  
Db 300 ACCGGTGAAACTCGTTTACAACCTGTTGAGAAAAATACTAAAATCATGAAAAAGTATCGT 359  
QY 342 CCCCAATGTTGCCAAATATGCTCTGATACCATCTGCTTATTCACCAATCCTGTGCA 401  
Db 360 AACTAGTGTCAATGATAGTGGCTTTGATGGTTTCTTCCCTAATTGCTGCAAAACCCAGTTGA 419  
QY 402 TGTTTTGACCTATATTAGCTATAAGGCGTCAAGGTTTCCACTAAGCAGAGTTATCGGCTC 461  
Db 420 TATCTTAACACGTTATGTTAAAGAAGTTACAGGTTTACAGCTGAACGTGTTATGGTTC 479  
QY 462 AGGTACAGTTCTGGATACCTGCTGTTTTAAATACATCTCGGAGAGCACCTTCAAGATCTC 521  
Db 480 TGGTACAGTGTCTGTGATAGTGAAGATTCAGATATTTAATAAGTAAAGATTTAGGTGTAC 539  
QY 522 ATCGGACAGCATCGATGCCGTGTGTAATTGGAGAACATGGTGATTCGGGTGTCCTGTCTG 581  
Db 540 ATCAAGTAGTGTTCACGCTAGCATTATAGGTGAACATGGTGACTCTGAACCTTGCACTTG 599  
QY 582 GTCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTTACTGCGAAAAAGCCAACACAT 641  
Db 600 GTCTCAAGCAAAACGTTGGAGGTATTTTCAGTGTATGATACATTTGAAAGAAGAACTGGTAG 659  
QY 642 ATTTGATCAGAATGCGTTCCATAGAATCTTTGAGCAAAACGCGAGAGCGCTGCTTACGATAT 701  
Db 660 CGATGCTAA-----AGCGAATGAAATTTATATTAATACAAGAGATGCTGCTTACGATAT 713  
QY 702 CATCAAGCGCAAAGGCTATCTTATGGAATCGCAGCGGGATTACTTCGCATAGTATAA 761  
Db 714 CATCAAGCTAAAGGATCTACGTATTATGGTATAGCTCTAGCACTATTACGTATTTCTTAA 773  
QY 762 GCGCATTTTAGAGGATACAGGATCCACACTTACAGTTTCAA---CCGTTGGTGATTATTT 818  
Db 774 AGCTTTACTAAATAATGAAATAAGTATTTTGACAGTTTCTAGTCAACTTAATGGTCAATA 833  
QY 819 TGGGGTTGAACAAAATGCTATAGCGTCCCTACCAAACTCAATAAAAAGTGGGGCTCATCA 878  
Db 834 TGGATTTAACGATGTTTATCTTGGCTTACCAACACTTATCAATCAAAATGGTGCACTTAA 893  
QY 879 AGTGGCTGAACCTTTTCACTCGATGAGAAGGAATAAGATTGATGGAATAATCAG 931  
Db 894 AATTATGAACACACCATTTAAATGATAACGAACACTACAATTACTAGAAAAATCAG 946

RESULT 6

AAH54991/c

ID AAH54991 standard; DNA; 3019 BP.

AC AAH54991;

XX 03-SEP-2001 (first entry)

XX S. epidermidis genomic polynucleotide sequence SEQ ID NO:4355.

XX Staphylococcus epidermidis SR1 strain; infection; diagnosis;

XX Staphylococcus epidermidis.











Db 6314 TTTGACTTATTCAACTTGGAAATCTCTGGCTCCCGAAAGAACGAGTAATCGTTTCAGG 6255  
Qy 465 TACAGTTCTGGATACTGCTCGTTTAAATAACATCCTCGGAGAGCACTTCAAGATCTCATC 524  
Db 6254 AACTTCACTAGATTCTGCTCGTTTCCGTCAGCAATGCCGAATTAGTTGACGTTGATGC 6195  
Qy 525 GGACAGCATCGATCGCTGTGTAATTGGAGAACATGGTGATTCGGGTGCTGCTGCTGTC 584  
Db 6194 ACGAAATGTCCATCGCTACATCTTAGGGGAACAGGAGATACAGAAATCCCGAGTTTGGTC 6135  
Qy 585 TCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTCGGAAAGCAACCAACACATATT 644  
Db 6134 ACATGCGAATGTCTGGCTTACAAAT---TTAGGAATGGTGAAATAATCCTGACGT 6078  
Qy 645 TGATCAGAAATCGTTCCATAGAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATATCAT 704  
Db 6077 CGATGAAGAAGCAATGGTTAATTTATTCTTCAACGTACGCGACGCTGCTTACACAATCAT 6018  
Qy 705 CAAGCGCAAGGCTATACCTTATATGGAATCGCAGCGGGATTACTTCGCATAGTAAAGGC 764  
Db 6017 CGAGAAAAAAGAGCTACTTTCTATGGAATCGCGGTTGCACTAGCGCTATCACTAAAGC 5958  
Qy 765 GATTTTAGAGGATACAGGATCCACACTTACAGTTTC---AACCGTTGCTGATTATTTGG 821  
Db 5957 TATCCTAAACGATGAAACTCTGTGTTACCATATCTGTTTATTAGAGGTGAATATGG 5898  
Qy 822 GTTGAACAATTTGCTATAAGCGTCCCTACCAACTCAATAAAAGTGGGCTCATCAAGT 881  
Db 5897 TCAAAACGATATTATATCGGTGCACCGCATCATCAACCGCAAGAGTTAAACAAGT 5838  
Qy 882 GGCTGAACCTTCACTCGATGAGAAGGAAATAGAAATGATGGAAAAATCAGCTAGTCAGAT 941  
Db 5837 CATTGAAATTCCTAATAACAGATGCTGAACAAGAAAAAATGGAAGCTTCTGCTTCTGATT 5778  
Qy 942 CAANTCAGTGATTGA 956  
Db 5777 AAAAGAAGTTATTGA 5763

RESULT 10  
AAV74735  
ID AAV74735 standard; DNA; 3264 BP.  
XX  
AC AAV74735;  
XX  
DT 16-MAR-1999 (first entry)  
XX  
DE Staphylococcus aureus contig SEQ ID #424.  
XX  
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;  
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
KW skin infection; surgical wound infection; scalded skin syndrome;  
KW toxic shock syndrome; ds.  
XX  
OS Staphylococcus aureus.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 1441..1500  
FT /\*tag= a  
FT /note= "these bases represent a line of missing text in  
FT the sequence listing in the specification. They  
FT are included to maintain the nucleotide numbering  
FT given in the specification for this DNA sequence"  
FT misc\_feature 3241..3264  
FT /\*tag= b  
FT /note= "these bases represent a line of missing text in  
FT the sequence listing in the specification. They  
FT are included to maintain the nucleotide numbering  
FT given in the specification for this DNA sequence"  
XX  
PN EP786519-A2.  
XX

PD 30-JUL-1997.  
XX  
PF 07-JAN-1997; 97EP-0100117.  
XX  
PR 05-JAN-1996; 96US-0009861.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Barash SC, Choi GH, Dillon PJ, Pannon MR, Kunsch CA;  
PI Rosen CA;  
XX  
DR WPI; 1997-374922/35.  
XX  
PT Polynucleotide(s) and proteins derived from Staphylococcus aureus  
PT stored on computer readable medium and used in the production of  
PT anti-S.aureus vaccines  
XX  
PS Claim 1; Page 1324-1325; 3271pp; English.  
XX  
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
CC of the invention. The DNA sequences are recorded on a computer readable  
CC medium, preferably selected from a floppy or hard disk, random access  
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
CC the S.aureus DNA sequences allows putative functions to be assigned so  
CC that protein-encoding or regulatory regions of commercial, therapeutic or  
CC industrial importance can be obtained. Specifically, sequences which are  
CC likely to encode antigens have been identified and these polypeptides can  
CC be used in a vaccine composition against S.aureus infection. The  
CC polypeptides can also be used in a kit for the immunodetection of  
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,  
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,  
CC skin and surgical wound infections, scalded skin syndrome, toxic shock  
CC syndrome, etc. Organisms transformed with the DNA sequences can be used  
CC for recombinant production of the polypeptides. The new DNA sequences  
CC (and their fragments) are useful as primers or probes for isolating  
CC homologues of any of the S.aureus DNA sequences contained on the  
CC computer readable medium.  
XX  
SQ Sequence 3264 BP; 1063 A; 462 C; 656 G; 996 T; 87 other;  
Query Match 17.8%; Score 172.6; DB 18; Length 3264;  
Best Local Similarity 52.2%; Pred. No. 1e-43;  
Matches 469; Conservative 0; Mismatches 409; Indels 21; Gaps 3;  
Qy 51 AAAAGTTGTGGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCT 110  
Db 412 AAAGTTGTATTAAATCGGAGATGGATCTGTAGGATCAAGCTATGCTTTGCAATGGTTAC 471  
Qy 111 CAGCGCATCGTTTCCGAGATTGTCCTTATCGACGTGAACAAAGACAAAGCAGAGGTGA 170  
Db 472 GCAAGGTGTGCTGATGAATTTGTAATTATTGACATTGC AAAAGACAAAGTAAAGCAGA 531  
Qy 171 AAGCATGGAACTTAAACCCAGCA-----CCTTCAAATACAAGGTCTCGAGCGG 221  
Db 532 TGTTCAAGATTAAACCATGGTACAGTCCACAGTCTTCCAGTTGATGTGAAAGCAGG 591  
Qy 222 TGATTATCCTGACTGCGCTGGCGGCGCCCATTTGTTATTGTCACATGTGGGATTAAACCAAAA 281  
Db 592 TGAATACGAAGACTGTAAAGATGCAGATTAGTTGTTATTACAGCTGGTGCACCTCAAAA 651  
Qy 282 AAATGGACAAACAAGGATGGATCTTGTCTGCAAAAATGCCAACATTATCTGGAAATCAT 341  
Db 652 GCCAGGTGAACACAGTTTACAATTAGTTGAAAAAATACTAAGATTATGAAGAGCATCGT 711  
Qy 342 CCCCATGTTGCCAAATATGCTCTCTGATACCATCCTGCTTATTGCCACGAATCCTGTGCA 401  
Db 712 TAAGAGTGTATGGATAGTGGCTTTGATGGATATTTCTTAATCGCGCAAAACCTGTAGA 771  
Qy 402 TGTGTTGACCTATATTAGCTATAAGCGCTCAGGGTTTCCACTAAGCAGAGTTATCGGCTC 461  
Db 772 CATTTTAAACAAGATTGTGTAAGAATATATCTGGATTACCAGCAGAGCGTGTATCGGTTTC 831  
Qy 462 AGGTACAGTTCTGGATACTGCTCGTTTAAATACATCCTCGGAGAGCACTTCAAGATCTC 521



Db 832 AGGTACTGTATTGGACAGTGCACGTTTACATATTTAATTAGCCAGAAGTGGTGGC 891  
QY 522 ATCGGACAGCATCGATGCCTGTGTAAATTGGAGAACATGGTGATTCGGGTGTGCTGTG 581  
Db 892 ACCTTCAAGTGTGACGCTAGTATTATTGGCGAGCATGGTGATCTGAACCTTGCAGTTG 951  
QY 582 GTCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAGCCAAACCAT 641  
Db 952 GTCACAAAGCAATGTAGCAGGTATTTTCAGTATATGACACATTAAGAA-----CAAAC 1005  
QY 642 ATTTGATCAGATGCGTTCATAGAACTTTTGAGCAAAACGCGAGACGCTGCTTACGATAT 701  
Db 1006 TGGTAGCGAAGCTAAAGCGGAAGAAATTTATGTGAATACACGTGACGCTGCTTATGAAAT 1065  
QY 702 CATCAAGCGCAAGGCTATCTTATATGGAATCGACGGGATTACTTCGCATAGTAAA 761  
Db 1066 TATCCAAGCTAAAGGTCACATACTATGTTATGATTAGCATTGATGCGCATTTCAA 1125  
QY 762 GCGGATTTTA-----GAGGATACAGGATCCACATTTACAGTTTCAACCGTTGGTGATTA 815  
Db 1126 AGCCATTTTAAATAATGAATAATGCTTTAAATGTTCTATACATATAGATGTTCAATA 1185  
QY 816 TTTTGGGTTGAACAAATTGCTATAAGCGTCCCTACCAACTCAATAAAGTGGGGCTCA 875  
Db 1186 TGGTGGTCAAAAGCGGTTTACCTAGTGTACCAACATTAGTTAAACCAACATGGCGCAGT 1245  
QY 876 TCAAGTGGCTGAACCTTTCACTCGATGAGAAAGGAATAGAAATTGATGGAAAATCAGCTA 934  
Db 1246 TAAAATTTATGAATGCCATTAAAGTGCCGAAGAACAAAGCGTTGTTTCGATAAATCTGTTA 1304

RESULT 11  
ID ABN68374 standard; DNA; 981 BP.  
XX.  
AC ABN68374;  
XX  
DT 01-JUL-2002 (first entry)  
XX  
DE Streptococcus polynucleotide SEQ ID NO 4661.  
XX  
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;  
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.  
XX  
OS Streptococcus pyogenes.  
XX  
PN WO200234771-A2.  
XX  
PD 02-MAY-2002.  
XX  
PF 29-OCT-2001; 2001WO-GH04789.  
XX  
PR 27-OCT-2000; 2000GB-0026333.  
PR 24-NOV-2000; 2000GB-0028727.  
PR 07-MAR-2001; 2001GB-0005640.  
XX  
PA (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Telford J, Massignani V, Margarit Ros VI, Grandi G, Fraser C;  
PI Tettelin H;  
XX  
DR WPI; 2002-352536/38.  
DR P-PSDB; ABP27743.  
XX  
PT New Streptococcus protein for the treatment or prevention of infection  
PT or disease caused by Streptococcus bacteria, such as meningitis, and  
PT for detecting a compound that binds to the protein -  
XX  
PS Claim 7; Page 3629; 4525pp; English.  
XX

CC The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins.  
XX  
SQ Sequence 981 BP; 278 A; 202 C; 199 G; 302 T; 0 other;  
Query Match 16.2%; Score 157.2; DB 24; Length 981;  
Best Local Similarity 50.7%; Pred. No. 4.2e-39;  
Matches 465; Conservative 0; Mismatches 438; Indels 15; Gaps 3;  
QY 51 AAAAGTTGTGGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCT 110  
Db 24 AAAAGTAATCTTGGTCGTGATGTCGCGTGGTTCATCATACGCTTTGTCATTGGTAAC 83  
QY 111 CAGCGGCATCGTTTCCGAGATTGTCCTTATCGCGTGAACAAAGACAGAGGGTGA 170  
Db 84 ACAAATATTGCTCAAGAACTTGGTATTATTGATATTTTAAAGAAAACCTCAAGGAGA 143  
QY 171 AAGCATGGACTTAAACCCAGCAGCACCT-----TCAATACAAGGTCTCGAGCGGGTGA 224  
Db 144 TGCAGAAAGATTAAAGCCACGCTTAGCTTTTACATCACCTAAAAAATCTATGCAGCTGA 203  
QY 225 TTATCCTGACTGCGTGGCGGGCCATTGTTATTGTACATGTGGGATTAAACCAAAAAA 284  
Db 204 CTACTCAGACTGTCAATGATGCTGATCTTGTGTTTAAACAGCTGGTGTCTCTCAAAACC 263  
QY 285 TGGACAAACAAAGGATGGATCTTGTGTCGCAAAAAAATGCCAACATATGCTGGAATCATCC 344  
Db 264 AGGTGAAACACAGCCTTGACCTTGTGTAAGAAAAAACCTTCGTATCAATAAAGAGGTTGTAC 323  
QY 345 CAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTGCCAGGAATCCTGTGATGT 404  
Db 324 ACAATCGTTGCATCTGTTTCAAGGAATCTTTCTTGTGCTGCTAACCCAGTTGACGT 383  
QY 405 TTTGACCTATATTAGCTATAAGCGGTGAGGTTTCCACTAAGCAGAGTTATCGGCTCAGG 464  
Db 384 TTTGACTTACTCAACATGGAATTTCTCAGGTTTCCCTTAAAGAGCGGTGTCATTTGTTT 443  
QY 465 TACAGTTCTGGATACCTGCTGTTTAAATACATCCTCGGAGAGCAGCTTCAAGATCTCATC 524  
Db 444 TACATCTCTAGATTACAGCCCGTTTCCGTCAGCTCTGGCTGCTAAATCGGTGTGATGC 503  
QY 525 GGACAGCATCGATGCTGTGTAATTGGAGAACATGGTGATTCCGGGTGTGCTGTGGTC 584  
Db 504 GCGTTGCGTTTACGCTTACATCATGCGGAGAACACGCGTGAATTTGCTGTTTGGTC 563  
QY 585 TCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCAAACCATATT 644  
Db 564 ACACGCTAACGTTGCTGGTGTAGGTCT--TTACGACTGGCTCCAAGCTAATCGTGACAT 620  
QY 645 TGATCAGAATCGGTTCCATAGAAATCTTTGAGCAAAACCGGAGAGCGCTGCTTACGATATCAT 704  
Db 621 TGATGACGAAGCCCTAGTTGATCTTTTTCATTTCTGTACGTGATGCTGCTTACTCAATCAT 680  
QY 705 CAAGCGCAAGGCTATACCTTCAATGGAATCGCAGCGGATTTACTTCGCATAGTAAAGGC 764  
Db 681 CAATAAAAAAGGTGCTACTTTCTACGGTATTGCAAGTTGCCCTAGCTCGTATTACAAAGC 740  
QY 765 GATTTTAGAGG-----ATACAGGATCCACACTTACAGTTTCAACCGTTGGTGATTATT 818

Db 741 TATTCCTGATGAAATGCTGTTCTTCGGCTTCTGTTTTCCAAAGGCCAATACGA 800  
QY 819 TGGGGTTGAACAATGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGCTCATCA 878  
Db 801 AGGTGTTGAAGATTGCTATATTGGTCAACAGCTATCGTTGGAGCGTATGGAATTGTTTCG 860  
QY 879 AGTGGCTGAACCTTCACTCGATGAGAAGGAAATAGAATTGATGGAAAAATCAGCTAGTCA 938  
Db 861 TCCAGTAAACATCCACTTAACGATGCTGAATTACAAAAATGCAAGCCCTCTGCAAAATCA 920  
QY 939 GATCAAAATCACTGATTGA 956  
Db 921 ATTGAAAGCTATCAATTGA 938

RESULT 12  
AAZ09139  
ID AAZ09139 standard; DNA; 945 BP.  
AC AAZ09139;  
XX  
DT 18-OCT-1999 (first entry)  
XX  
DE B. flavum lactate dehydrogenase DNA.  
XX  
KW Lactate dehydrogenase; destruction; lactic acid; ss.  
XX  
OS Brevibacterium flavum.

Key Location/Qualifiers  
FH 1..945  
FT /\*tag= a  
FT /product= "lactate dehydrogenase"

PN JP11206385-A.  
XX  
PD 03-AUG-1999.  
XX  
PF 28-JAN-1998; 98JP-0030594.  
XX  
PR 28-JAN-1998; 98JP-0030594.  
XX  
PA (MITU ) MITSUBISHI CHEM CORP.  
XX  
DR WPI; 1999-486360/41.  
DR P-PSDB; AAY25997.

PT New lactate dehydrogenase gene and strain destroying gene - useful  
PT for formation of lactic acid  
XX  
PS Claim 2; Page 10-11; 13pp; Japanese.

CC This invention describes the isolation of a novel lactate dehydrogenase  
CC protein from Brevibacterium flavum strain MJ-233. The invention also  
CC describes a microbial strain destroying the lactate dehydrogenase gene of  
CC a microbe in which the lactate dehydrogenase gene is destroyed by a  
CC homologous recombination of the above DNA or the above recombinant vector  
CC DNA with the lactate dehydrogenase gene on the chromosome DNA of the  
CC microbe cell and a method for the preparation of an amino acid or an  
CC organic acid in which the above strain destroying the lactate  
CC dehydrogenase gene is cultured in a medium and the amino acid or the  
CC organic acid (except lactic acid) is collected from the culture.  
CC Formation of lactic acid in the preparation of an amino acid and an  
CC organic acid can be decreased with no control of oxygen concentration  
CC during culture. This sequence encodes the lactate dehydrogenase protein  
CC described in the specification.

XX  
SQ Sequence 945 BP; 238 A; 285 C; 237 G; 185 T; 0 other;

Query Match 16.0%; Score 155.8; DB 20; Length 945;  
Best Local Similarity 50.9%; Pred. No. 1.1e-38;  
Matches 457; Conservative 0; Mismatches 422; Indels 18; Gaps 3;

QY 52 AAAGTTGTGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC 111  
Db 22 AAGATTGTCTTATTGGCGCAGAGATGTTGGAGTTGCATACGCATACGCTGATCAAC 81  
QY 112 AGCGGATCGTTTTCCGAGATTGCTCTTATCGACGTGAACAAGACAAAGCAGAGGGTGAA 171  
Db 82 CAGGGCATGGCAGATCACCTTGGCATCATCGACATCGATGAAGAAGAACTCGAAGGCAAC 141  
QY 172 AGCATGGACTTAAACCAACGCA-----GCACCTTCAAATACAAGGTCTCGAGCGGGT 222  
Db 142 GTCATGGACTTAAACCATGTTGTGTGGGCCGATTCCCGCACCCGCTACCAAGGGC 201  
QY 223 GATTATCCTGACTGCGTGGCGGCCCATTTGTTATTGTACATGTGGGATTAAACCAAAA 282  
Db 202 ACCTACGCTGACTGCGAGGACGAGCCCATGTTGTCAATTTGTGCGGCGCAGCCCAAG 261  
QY 283 AATGACAAAAAAGGATGGATCTTGTGCTGCAAAAAATGCCAACAATATGCTGGAATCATC 342  
Db 262 CCAGCGAAAACTCGCCTCCAGCTGGTGGACAAAAACGTCAGGATTATGAAGTCCATCGTT 321  
QY 343 CCCAATGTTGCCAAATATGCTCCTGATACCATCTGCTTATTGCCACGAATCCTGTCGAT 402  
Db 322 GCGATGTCATGGCCAGCGGATTCGACGGCATCTTCCTCGTAGCCTCCAACCCAGTGGAT 381  
QY 403 GTTTTGACCTATATTAGCTATAAGGCGTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCA 462  
Db 382 ATCCTCACCTACGAGTGTGGAATTTCTCGGCCCTGGAATGGAACCGCGTATCGGCTCC 441  
QY 463 GGTACAGTTCTGGATACTGCTCGTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCA 522  
Db 442 GGAATGTCCTGGACTCCGCTAGATTCCGCTACATGCTCGGGAACCTATGAAGTGGCA 501  
QY 523 TCGGACAGCATCGATGCTGTGTAATTTGGAGAACATGTTGCGGTGTCCTGCTGCG 582  
Db 502 CCAAGCTCCGTCACGCTACATCATCGGCGAACACGCGGACACTGAACCTCCAGTCCCTG 561  
QY 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCCAACCATATA 642  
Db 562 TCCTCGCGGACCATCGCAGGCGTATCGCTTAG-----CCGATGCTAGACAAAGACCCA 615  
QY 643 TTTGATCAGAAATGCGTTCCATAGAATCTTTGAGCAAACGCGAGACGCTGCTTACGATATC 702  
Db 616 GAGCTTGAGGGCGCTCTAGAGAAAAATTTTCGAAGACACCCGCGACGCCCTATCACATC 675  
QY 703 ATCAAGCGCAAAGGCTATATCTTCATATGGAATCGCAGCGGGAATTACTTCGCATAGTAAAG 762  
Db 676 ATCGACGCCAAGGGCTCCACTTCTACGGCATCGGCATCGGCTTGTGCTGCATCACCCGC 735  
QY 763 GCGATTTTAGAGGATACAGGATCCACACTTACAGT---TTCAACCGTTGGTGAATATTTT 819  
Db 736 GCAATCCTACAAAACCAAGACGTTGCAGTCCAGTCTCTGCATGCTCCACGGTGAATAC 795  
QY 820 GGGTTGAACAAATTTGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGCTCATCAA 879  
Db 796 GGTGAGGAAGACATCTACATCGGCACCCCGCAGTAGTAACCCCGGAGGCATCCGCCGC 855  
QY 880 GTGCTGAACCTTTCACTCGATGAGAAGAAATAGAATTGATGAAAAAATCAGCTAGT 936  
Db 856 GTTGTGAACTAGAAATCACGGACCATGAGATGGAACGCTTCAAGCATTCGCGCAAT 912

RESULT 13  
AAAF71424  
ID AAF71424 standard; DNA; 1065 BP.  
XX  
AC AAF71424;  
XX  
DT 30-APR-2001 (first entry)  
XX

DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:129.  
XX  
KW Corynebacterium glutamicum; carbon metabolism and energy production;

KW SMP protein; sugar metabolism and oxidative phosphorylation protein;  
KW fine chemical production; organic acid; proteinogenic amino acid;  
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;  
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;  
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;  
KW diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.  
XX  
OS Corynebacterium glutamicum.  
XX  
PN WO200100844-A2.  
XX  
PD 04-JAN-2001.  
XX  
PF 23-JUN-2000; 2000WO-IB00943.  
XX  
PR 25-JUN-1999; 99US-0141031.  
PR 08-JUL-1999; 99DE-1031412.  
PR 08-JUL-1999; 99DE-1031413.  
PR 08-JUL-1999; 99DE-1031419.  
PR 08-JUL-1999; 99DE-1031420.  
PR 08-JUL-1999; 99DE-1031424.  
PR 08-JUL-1999; 99DE-1031428.  
PR 08-JUL-1999; 99DE-1031431.  
PR 08-JUL-1999; 99DE-1031433.  
PR 08-JUL-1999; 99DE-1031434.  
PR 08-JUL-1999; 99DE-1031510.  
PR 08-JUL-1999; 99DE-1031562.  
PR 08-JUL-1999; 99DE-1031563.  
PR 08-JUL-1999; 99DE-1031634.  
PR 09-JUL-1999; 99DE-1032180.  
PR 09-JUL-1999; 99DE-1032227.  
PR 09-JUL-1999; 99DE-1032230.  
PR 09-JUL-1999; 99US-0143208.  
PR 14-JUL-1999; 99DE-1032924.  
PR 14-JUL-1999; 99DE-1032973.  
PR 14-JUL-1999; 99DE-1033005.  
PR 27-AUG-1999; 99US-0151572.  
PR 31-AUG-1999; 99US-0151572.  
PR 03-SEP-1999; 99DE-1042076.  
PR 03-SEP-1999; 99DE-1042079.  
PR 03-SEP-1999; 99DE-1042086.  
PR 03-SEP-1999; 99DE-1042087.  
PR 03-SEP-1999; 99DE-1042088.  
PR 03-SEP-1999; 99DE-1042095.  
PR 03-SEP-1999; 99DE-1042123.  
PR 03-SEP-1999; 99DE-1042125.  
XX  
PA (BADI ) BASF AG.  
XX  
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
XX  
XX WPI; 2001-061975/07.  
DR P-PSDB; AAB79307.  
XX  
PT New isolated Corynebacterium glutamicum nucleic acid encoding a sugar  
PT metabolism and oxidative phosphorylation protein for production or  
PT modulation of production of fine chemicals e.g. amino acids,  
PT carbohydrates or enzymes -  
XX  
PS Claim 3; Page 332-333; 1246pp; English.  
XX  
CC AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar  
CC metabolism and oxidative phosphorylation (SMP) proteins given in  
CC AAB79243 to AAB 79633 which are involved in carbon metabolism and  
CC energy production. The C. glutamicum SMP gene can be used in vectors  
CC (II) for expression in host cells and production or modulation of  
CC production of fine chemicals, such as, an organic acid, a proteinogenic  
CC or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,  
CC a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty  
CC acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a  
CC cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins  
CC (III) encoded by them are used for diagnosing the presence or activity of  
CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells  
CC containing them are used to map genomes of organisms related to

CC C. glutamicum, identify and localise C. glutamicum sequences of interest,  
CC in evolutionary studies, in determining SMP protein regions required  
CC for function, in modulating SMP protein activity, in modulating the  
CC metabolism of sugars, and in modulating high-energy molecule production  
CC in a cell (i.e. ATP, NADPH).  
XX  
SQ Sequence 1065 BP; 264 A; 304 C; 279 G; 218 T; 0 other;  
  
Query Match 16.0%; Score 155.8; DB 22; Length 1065;  
Best Local Similarity 50.9%; Pred. No. 1.2e-38;  
Matches 457; Conservative 0; Mismatches 422; Indels 18; Gaps 3;  
  
QY 52 AAAGTTGTGGTAGTGGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC 111  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
122 AAGATTGTCTCATTTGGCGCAGGAGATGTTGGAGTTGCATACGCATACGACTGATCAAC 181  
QY 112 AGCGGCATCGTTTCCGAGATTGTCCTTATCGACGTGAACAAAGACAAAGCAGAGGGTGAA 171  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
182 CAGGCGATGGCAGATCACCTTGGCGATCATCGACATCGATGAAAGAAACTCGAAGGCAAC 241  
QY 172 AGCATGGACTTAAACCAACGCA-----GCACCTTCAATACAAAGGTCTCGAGCGGGT 222  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
242 GTCATGGACTTAAACCATGTTGTGTGGCCGATTCGCCACCCCGCTCACCAAGGCG 301  
QY 223 GATTATCTGACTGCGCTGGCGGCCCATTTGTTATTGTCACATGTGGGATTAAACCAAAA 282  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
302 ACCTACGCTGACTGCGAAGACGAGCCATGTTGTTCATTTGTCCGCGCAGCCCAAG 361  
QY 283 AATGACAACAAAGGATGGATCTTGCTGCAAAAAATGCCAACATTATGCTGGAATCATC 342  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
362 CCAGCGAGACCCGCTCCAGCTGGTGGACAAAAACGTCAAGATTATGAAATCCATCGTC 421  
QY 343 CCCAATGTTGCCAAATATGCTCTCTGATACCATCTCTGTTATTGCCACGAATCCTGTGAT 402  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
422 GCGATGTCATGGACAGCGGATTTCGACGGCATCTTCTCTGTCGCTCCAACCCAGTGGAT 481  
QY 403 GTTTTGACCTATATTAGCTATAAGCGCTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCA 462  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
482 ATCCTGACCTACGCAGTGTGAAATTTCTCGGCTTGAATGGAACCGCGTGATCGGCTCC 541  
QY 463 GGTACAGTTCTGGATACCTGCTCGTTTAAATACATCTCGGACAGCACTTCAAGATCTCA 522  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
542 GGAAGTCTCTGGACTCGCTCGCTCGATTCCGCTACATGCTGGCGAACTCTACGAAGTGCA 601  
QY 523 TCGGACAGCATCGATGCTCTGTGTAATTGGAGAACATGGTGATTCCGGTGTGCTGTGG 582  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
602 CCAAGCTCCGTCCACGCCTACATCATCGGCGAACACGGCGACACTGAACCTCCAGTCTG 661  
QY 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGATTACTGCGAAAAAGCCCAACACATA 642  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
662 TCCTCCGGACCATCGCAGGCGTATCGCTTAGCCGAATGCTGGACAAAGACCCAGAGCT- 720  
QY 643 TTGATCAGAATGCGTTCCATAGAAATCTTTGAGCAACCGCAGACGCTGCTTACGATATC 702  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
721 -----TGAGGGCGCTCTAGAGAAAAATTTTCGAAGACACCCCGCAGCTGCCTATCACATT 775  
QY 703 ATCAAGCGCAAAAGGCTATACTTCAATATGGAATCGCAGCGGATTACTTCGCATAGTAAAG 762  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
776 ATCGACGCAAGGGCTCCACTTCTTACGGCATCGGCTGTTGCTTCGCATCACCCGC 835  
QY 763 GCGATTTTAGAGGATACAGGATCCACACTTACAGT---TTCAACCGTTGGTGATTATTTT 819  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
836 GCAATCTTACAGAACCAAGACGTTGCACTCCAGTCTCTGCACTGCTCCACGGTGAATAC 895  
QY 820 GGGTTGAACAAATTGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGCTCATCA 879  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
896 GGTGAGGAAGACATCTACATCGGCACCCAGCTGTGGTGAACCCGCGAGGCATCCGCCGC 955  
QY 880 GTGGCTGAATTTCACTCGATGAGAGGAAATAGAAATTGATGGAATAATCAGCTAGT 936  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
956 GTTGTCGAACTAGAAATCACCGACACGAGATGGAACGCTTCAAGCATTCGCGCAAT 1012



```

RESULT 14
AAH68175
ID    AAH68175 standard; DNA; 942 BP.
XX
AC    AAH68175;
XX
DT    26-SEP-2001 (first entry)
DE
DE    C glutamicum coding sequence fragment SEQ ID NO: 3210.
XX
KW    Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW    organic acid synthesis; ds.
XX
OS    Corynebacterium glutamicum.
XX
PN    EP1108790-A2.
XX
PD    20-JUN-2001.
XX
PF    18-DEC-2000; 2000EP-0127688.
XX
PR    16-DEC-1999; 99JP-0377484.
PR    07-APR-2000; 2000JP-0159162.
PR    03-AUG-2000; 2000JP-0280988.
XX
PA    (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI    Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI    Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR    WPI; 2001-376931/40.
DR    P-PSDB; AAG92956.
XX
PT    Novel polynucleotides derived from Coryneform bacteria, for identifying
PT    mutation point of a gene, measuring expression of a gene, analysing
PT    expression profile or pattern of a gene and identifying homologous gene
PT
PS    Claim 8; SEQ ID NO: 3210; 246pp + Sequence Listing; English.
XX
CC    The present invention provides a number of nucleotide and protein
CC    sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC    are useful for identifying the mutation point of a gene derived from a
CC    mutant of coryneform bacterium, measuring expression amount and
CC    analysing the expression profile or expression pattern of a gene derived
CC    from Coryneform bacterium, and identifying a homologue of a gene derived
CC    from coryneform bacterium. Coryneform bacteria are useful for producing
CC    amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC    particularly L-lysine. The present sequence is a nucleic acid described
CC    in the exemplification of the invention.
CC    Note: The sequence data for this patent did not form part of the printed
CC    specification, but was obtained in electronic format directly from the
CC    European Patent Office.
XX
SQ    Sequence 942 BP; 232 A; 284 C; 243 G; 183 T; 0 other;

Query Match      15.9%; Score 154.2; DB 22; Length 942;
Best Local Similarity 50.8%; Pred. No. 3.6e-38;
Matches 456; Conservative 0; Mismatches 423; Indels 18; Gaps 3;

QY    52 AAAGTTGTGTAGTGGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC 111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    22 AAGATTGTCTCTCATTTGGCGCAGGACATGTTGGAGTTGCATACGCATACGCACTGATCAAC 81
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY    112 AGCGGCATCGTTTCCGAGATTGTCTTATCGACGTGAACAAGACAAAGCAGAGGTGAA 171
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    82 CAGGGCATGGCAGATCACCTTGGCATCGACATCGATGATGAAGAACTCGAAGGCAAC 141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY    172 AGCATGGACTTAAACACCGCA-----GCACCTTCAATACAAGGTCTCGAGCGGGT 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    142 GTCATGGACTTAAACCATGTGTGTGTGGCCGATTCCCGACCCCGGTACCAAGGGC 201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY    223 GATTATCTGTACTGCGTGGCGCGGCCCATTTGTTATTGTACATGTGGGATTAAACCAAAA 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



PA (KYOW ) KYOWA HAKKO KOGYO KK.

XX  
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI; 2001-376931/40.

Novel polynucleotides derived from Coryneform bacteria, for identifying mutation point of a gene, measuring expression of a gene, analysing expression profile or pattern of a gene and identifying homologous gene

PS Disclosure; SEQ ID NO: 7069; 246pp + Sequence Listing; English.

The present invention provides a number of nucleotide and protein sequences from the *Corynebacterium* bacterium *Corynebacterium glutamicum*. These are useful for identifying the mutation point of a gene derived from a mutant of *corynebacterium* bacterium, measuring expression amount and analysing the expression profile or expression pattern of a gene derived from *Corynebacterium* bacterium, and identifying a homologue of a gene derived from *corynebacterium* bacterium. *Corynebacterium* bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a nucleic acid described in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC European Patent Office.

XX  
SQ Sequence 309400 BP; 70133 A; 86477 C; 83115 G; 69675 T; 0 other;

**Query Match**      **15.9%**;    **Score 154.2**;    **DB 22**;    **Length 309400**;

Best Local Similarity 50.8%; Pred. No. 8.1e-37;

Matches 456; Conservative 0; Mismatches 423; Indels 18; Gaps 3;

QY	52	AAAGTTGTGGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC	111
Db	113369	AAGATTGTCTCATTTGGCGCAGGAGATGTTGGAGTTGCATACGCATACGCATGATCAAC	113310
QY	112	AGCGGCACTGTTTCCGAGATTGTCTTATCGACGTGAACAAAGACAAAGCAGAGGTTGAA	171
Db	113309	CAGGGCATGGCAGATCACCTTGGCATCATCGACATCGATGAAAGAAACTCGAAGGCAAC	113250
QY	172	AGCATGGACTTAAACCAAGCA-----GCACCTTCAAATACAAAGTCTCGAGCGGCT	222
Db	113249	GTCATGGACTTAAACCATGGTGTGTGTGGCCGATCCCGCACCCGCGTCACCAAGGC	113190
QY	223	GATTATCTGACTGGCTGGCGGGCCATTGTTATTGTACATGTGGGATTAAACCAAAA	282
Db	113189	ACCTACGTGACTGCGAAGACGAGCCATGGTTGTCAATTTGTGCCGGCGCAGCCCAAAG	113130
QY	283	AATGGACAAAACAGGATGGATCTTGCTGCAAAAATGCCAACATTATGCTGGAAATCATC	342
Db	113129	CCAGGCGAGACCCGCCCTCCAGCTGGTGGACAAAACGTCAAGATTATGAAATCCATCGTC	113070
QY	343	CCCAATGTTGCCAAATATGCTCTGATACCATCCCTGCTTATTGCCACGAATCCTGTGAT	402
Db	113069	GGCGATGTCATGGACAGCGGATTGACCGGCATCTTCTCTGTGGCGTCCAACCCAGTGGAT	113010
QY	403	GTTTTGACCTATATTAGCTATAAGCGGTACGGGTTTCCACTAAGCAGATTATCGGCTCA	462
Db	113009	ATCCTGACCTACGCAGTGTGAAAATTCTCCGGCTTGAATGGAACCGCGTGATCGGCTCC	112950
QY	463	GGTACAGTTCTGGATACTGCTCGTTTAAATACATCTCGGAGAGCACTTCAAGATCTCA	522
Db	112949	GGAAC TGTCCTGGACTCCGCTCGATTCCGCTACATGCTGGCGGAACCTACGAAGTGGCA	112890
QY	523	TCGGACAGCATCGATGCCTGTGTAATTGGAGAACATGGTGATTTCGGGTGTGCCTGTCTGG	582
Db	112889	CCAAGCTCCGTCCACGCCTACATCATCGCGGAACACGCGCACACTGAACTTCCAGTCCGTG	112830
QY	583	TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTCGGAAAAAGCCAAACCATATA	642
Db	112829	TCCTCCGGACCATCGCAGGCGTATCGCTTAGCCGAATGCTGGACAAAGACCCAGAGCT-	112771

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 08:56:31 ; Search time 62 Seconds  
(without alignments)  
4807.900 Million cell updates/sec

Title: US-09-992-430B-21  
Perfect score: 972  
Sequence: 1 atgtccaagatacaaaagtc.....ttgagcatctggagatcaat 972

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Query		DB	ID	Description
			Length				
1	183.8	18.9	993	4	US-09-134-001C-1491	Sequence 1491, Ap	
C 2	153.6	15.8	5278	4	US-08-961-527-227	Sequence 227, App	
C 3	147.2	15.1	5449	4	US-09-546-990-1	Sequence 1, Appli	
4	125	12.9	1661	4	US-09-166-350-31	Sequence 31, Appl	
5	109.8	11.3	1260	4	US-09-166-350-22	Sequence 22, Appl	
6	101.2	10.4	1240	1	US-08-869-506-1	Sequence 1, Appli	
7	101.2	10.4	1240	3	US-09-128-967-1	Sequence 1, Appli	
8	93.8	9.7	1116	4	US-09-535-381-1	Sequence 1, Appli	
9	93.8	9.7	6072	4	US-09-535-381-3	Sequence 3, Appli	
10	89.2	9.2	1912	1	US-08-270-013B-1	Sequence 1, Appli	
11	89.2	9.2	1912	1	US-08-838-418-1	Sequence 1, Appli	
12	81	8.3	978	4	US-09-134-001C-2696	Sequence 2696, App	
13	79.8	8.2	566	4	US-09-328-111-455	Sequence 455, App	
14	74.2	7.6	5024	1	US-08-920-812-7	Sequence 7, Appli	
15	74.2	7.6	5024	1	US-08-920-827-7	Sequence 7, Appli	
16	74.2	7.6	5024	1	US-08-921-177-7	Sequence 7, Appli	
17	74.2	7.6	5024	1	US-08-362-577C-7	Sequence 7, Appli	
18	74.2	7.6	5024	2	US-08-920-828-7	Sequence 7, Appli	
C 19	62.2	6.4	519	4	US-09-222-575-160	Sequence 160, App	
20	59	6.1	1679	3	US-08-676-882-1	Sequence 1, Appli	
21	58.4	6.0	951	5	PCT-US94-03796-1	Sequence 1, Appli	
22	52.4	5.4	497	4	US-09-370-838-129	Sequence 129, App	
23	51.4	5.3	493	4	US-09-370-838-181	Sequence 181, App	
24	36.2	3.7	429	4	US-09-071-035-89	Sequence 89, Appl	
C 25	36	3.7	87350	3	US-08-781-891-79	Sequence 79, Appl	
C 26	36	3.7	87543	4	US-09-791-211-3	Sequence 3, Appli	
27	35.4	3.6	259	4	US-09-071-035-91	Sequence 91, Appl	

28	34.4	3.5	471	4	US-09-134-001C-1397	Sequence 1397, Ap
C	29	34	7218	1	US-08-232-463-14	Sequence 14, Appl
30	33.8	3.5	608	4	US-09-328-111-532	Sequence 532, App
31	33.8	3.5	619	4	US-09-328-111-400	Sequence 400, App
C	32	33.8	2912	4	US-09-307-143-3	Sequence 3, Appli
33	33.4	3.4	477	1	US-08-403-866-13	Sequence 13, Appl
34	33.4	3.4	2231	1	US-08-403-866-14	Sequence 14, Appl
35	33.4	3.4	12720	1	US-08-403-866-11	Sequence 11, Appl
36	32.8	3.4	3292	4	US-09-336-447A-14	Sequence 14, Appl
37	32.8	3.4	19227	3	US-09-090-793-13	Sequence 13, Appl
38	32.8	3.4	40138	3	US-09-090-793-12	Sequence 12, Appl
39	32.2	3.3	1203	1	US-08-602-010A-7	Sequence 7, Appli
40	32.2	3.3	1203	3	US-08-680-726A-7	Sequence 7, Appli
41	32.2	3.3	1203	3	US-09-092-409-7	Sequence 7, Appli
42	32.2	3.3	5495	1	US-08-602-010A-1	Sequence 1, Appli
C	43	32.2	5495	1	US-08-602-010A-2	Sequence 2, Appli
44	32.2	3.3	5495	1	US-08-680-726A-1	Sequence 1, Appli
C	45	32.2	5495	1	US-08-680-726A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-134-001C-1491  
; Sequence 1491, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 1491  
; LENGTH: 993  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-1491

Query Match	18.9%	Score 183.8;	DB 4;	Length 993;
Best Local Similarity	53.0%	Pred. No. 2.2e-49;		
Matches	473;	Conservative	0;	Mismatches 402; Indels 18; Gaps 3;
Qy	51	AAAAGTTGTGGTAGTGGGAGTGGGAAAGTGTGGGTCTGCCACAGCGTATACGTTCTTCT	110	
Db	60	AAAAGTTGTTTAGTAGGAGACGGTCCGTAGGTTCAAGTTATGTCATTTGCTATGGTGAC	119	
Qy	111	CAGCGCATCGTTTCCGAGATTTCCTTATCGACGTGAACAAAGACAGAGGGTGA	170	
Db	120	TCAAGGAATGCAGATGAATTTGTAATTATGATATTGCAAAAGATAAAGTGAAGCAGA	179	
Qy	171	AAGCATGGACTTAAACCACGACGACCTTCAAATACAAGGTCTCG-----AGCGGG	221	
Db	180	CGTTAAAGATTAAACCATGGTGCACTTTACAGTTCTTCCACGAGTACTGTAAAGCTGG	239	
Qy	222	TGATTATCCTGACTGGCTGGCGGCCCATTTGTTATTGTGCATGTGGGATTAACCAAAA	281	
Db	240	AGATATGAAGATTGTAAGATGCAGATTTAGTTGTTATTACAGCAGGTGCACCTCAAAA	299	
Qy	282	AAATGGACAAACAAGGATGGATCTTGTCTGCAAAAAATGCCAAATATGCTGGAAATCAT	341	
Db	300	ACCGGGTGAAGAACTCGTTTACAACCTGTTTGAGAAAAATACTAAAATCATGAAAAGTATCGT	359	
Qy	342	CCCCAATGTTGCCAAATATGCTCTGTATACCATCCTGCTTATTGCCACGAATCCTGTGCA	401	
Db	360	AACTAGTGTGATGGTAGTGGCTTTGATGGTTTCTTCCTTAATTGCTGCAAAACCAGTTGA	419	

QY	402	TGTTTGGACCTATATTAGCTATAAAGCGCTCAGGGTTTCCACTAAGCAGAGTTATCGGCTC	461
DB	420	TATCTTAACACCGTTATGTTAAAGAAGTTACAGGTTTACCAGCTGAACGTGTTATTGGTTC	479
QY	462	AGGTACAGTTCTGGATACTGCTCGTTTAAATAACATCCTCGGAGAGCACTTCAAGATCTC	521
DB	480	TGGTACAGTGGTTGATAGTGCAAGATTCAGATATTTAATAAGTAAAGAAATTAGGTGTTAC	539
QY	522	ATCGGACAGCATCGATGCCGTGTGTAAATTGGAGAACATGGTGATTCGGGTGTGCCCTGTCTG	581
DB	540	ATCAAGTAGTGTTCACGCTAGCAATTATAGGTGAACATGGTGACTCTGAACTTGCAGTTG	599
QY	582	GTCCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCAAACCAT	641
DB	600	GTCCTCAAGCAAACGTTGGAGGTATTTCAAGTGTATGATACATTTGAAAGAAGAACTGGTAG	659
QY	642	ATTTGATCAGAATCGGTTCCATAGAAATCTTTTGAGCAACCGCAGACGCTGCTTACGATAT	701
DB	660	CGATGCTAA-----AGCGAATGAATTTATATTAATACAAGAGATGCTGCTTACGATAT	713
QY	702	CATCAAGCGCAAGGCTATACTTCATATGGAATCGCAGCGGATTACTTCGCATAGTAAA	761
DB	714	CATTCAAGCTAAAGGATCTACGTATTATGGTATAGCTCTAGCACTATTACGTAATTTCTAA	773
QY	762	GGCGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAA--CCGTTGGTGATTATTT	818
DB	774	AGCTTTACTAAATAATGAAAAATAGTATTTTGACAGTTTCTAGTCAACTTAATGGTCAATA	833
QY	819	TGGGGTTGAACAAATTGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGGCTCATCA	878
DB	834	TGGATTTAACGATGTTTTATCTTGGCTTACCAACACTTATCAATCAAAATGGTGCAGTTAA	893
QY	879	AGTGGCTGAATTTCACTCGATGAGAGGAAATAGAAATTGATCGAAAAATCAG	931
DB	894	AATTTATGAACACCATTAATGATAACGAACTACNAATTACTAGAAAAATCAG	946

## RESULT 2

US-08-961-527-227/c  
; Sequence 227, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 227:  
; SEQUENCE CHARACTERISTICS:





Db 401 GAGGGAGAAAGCCGCTCTTAATTGGTCCAGCGTAACGTGAACATATTTAAATTCATCATT 460  
Qy 343 CCCAATGTTGCCAATATATGCTCCTGTATACCATCCTGCTTATTGCCAGGAATCCTGTGAT 402  
Db 461 CCTAATGTTGTAATAATACAGCCCGAACTGCAAGTGTCTTATTGTTTCAAATCCAGTGGAT 520  
Qy 403 GTTTTGACCTATATTAGCTATAAGGCGTCAAGGCTTCCACTAAGCAGAGTTATCGGCTCA 462  
Db 521 ATCTTGACCTACGTGGCTTGAAGATAAGTGGTTTTCCCAAAACCGTGTATTGGAAGT 580  
Qy 463 GGTACAGTTCTGGATACCTGCTCGTTTTTAATAATGGAACATGCTCGGAGAGCACTTCAAGATCTCA 522  
Db 581 GGTGCAATCTGGATTCAGCCCGATTCCGTTACCTGATGGGGAAAGGCTGGGAGTTTCA 640  
Qy 523 TCGACAGCATCGATGCTGTGTAATTGGAACATGCTGATTCGGGTGCTGCTGCTG 582  
Db 641 CCATTAAGCTGTATGGTGGTCCCTTGGGAAACATGGAGATTCCAGTGTGCTGTATGG 700  
Qy 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGATTACTCGGAAAGCAACCAACATA 642  
Db 701 AGTGAATGAATGCTGGTGTCTCTCTGAAGACTCTGCACCCAGATTAGGGACTGAT 760  
Qy 643 TTTGATCAGAATGCTTCCATAGAACTTTTGACAAACGCGAGACGCTGCTTACGATATC 702  
Db 761 AAAGATAAGGAACAGTGGAAAGAGGTTCCACAGCAGGTGGTTGAGAGTGTCTTATGAGGTG 820  
Qy 703 ATCAAGCGCAAAGGCTATATTTCATATGGAATCGCAGCGGGATTACTTCCGATAGTAAAG 762  
Db 821 ATCAAACTCAAAGGCTACACATCTCTGGGCTATTGGAATCTCTGTAGCAGATTGGCAGAG 880  
Qy 763 GCGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAACCGTTGGTGATTTT 819  
Db 881 AGTATAATGAAGAACTTAGGCGGTGACCCAGTTTCCACCATGATTAAAGGCTCTT 937

RESULT 5

US-09-166-350-22  
; Sequence 22, Application US/09166350A  
; Patent No. 6440663  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew  
; APPLICANT: Chen, Yao  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Jager, Lloyd  
; APPLICANT: Knuth, Alex  
; TITLE OF INVENTION: Renal Cancer Associated Antigens and  
; TITLE OF INVENTION: Uses Therefor  
; FILE REFERENCE: L0461/7051  
; CURRENT APPLICATION NUMBER: US/09/166,350A  
; CURRENT FILING DATE: 1998-10-05  
; EARLIER APPLICATION NUMBER: US 09/166,350  
; EARLIER FILING DATE: 1998-10-05  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 22  
; LENGTH: 1260  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-166-350-22

Query Match 11.3%; Score 109.8; DB 4; Length 1260;  
Best Local Similarity 47.7%; Pred. No. 1.9e-25;  
Matches 361; Conservative 0; Mismatches 387; Indels 9; Gaps 1;  
Qy 61 GTAGTGGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTCAGCGGCATC 120  
Db 130 GTAGTGGGTGTTGGACAAAGTGGTATGGCGTGTGTATCAGCATTTCTGGGAAAGTCTCTG 189  
Qy 121 GTTCCGAGATTGCTCTTATCGACGTGAACAAAGCAAAAGCAGAGGTTGAAAGCATGGAC 180  
Db 190 GCTGATGAACCTGCTCTTGTGGATGTTTGGGAAGATAAGCTTAAAGGAGAAATGATGAT 249

Qy 181 TTAACCACAG-----CAGCACCTTCAAATACAGGTCTCGAGCGGTGATTATCT 231  
Db 250 CTGCAGCATGGGAGCTTATTCTTTCAGACACCTAAATTTGTGGCAGATAAAGATTATCT 309  
Qy 232 GACTGCGCTGGCGCGCCATTGTTATTGTACATGTGGGATTAACCAAAAAATGGACAA 291  
Db 310 GTGACCGCCAATTCTAAGATTGTAGTGTAACTGCAGAGTCCGTGAGCAAGAGGGGAG 369  
Qy 292 ACAAGGATGGATCTTGTCTGCAAAAAATGCCAACATTTATCTGGAAATCATCCCCATGTT 351  
Db 370 AGTCGGCTCAATCTGGTGACAGAAATGTTAATGTCTTCAAATTCATTATTCCTCAGATC 429  
Qy 352 GCCAAATATGCTCCTGTATACCATCCTGCTTATTGCCCACGAATCCTGTGATGTTTGACC 411  
Db 430 GTCAAGTACAGTCTGTGATTCATATAATTTGTGTTTCCAACCCAGTGGACATCTTTAG 489  
Qy 412 TATATTAGCTATAAGGCGTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCAGGTACAGTT 471  
Db 490 TATGTTACCTGGAACTAAGTGGATTACCCAAACACCGCTGATTGGAAGTGGATGTAAT 549  
Qy 472 CTGATACTGCTCGTTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCATCGGACAGC 531  
Db 550 CTGGATTCTGTAGATTTCGCTACCTTATGGCTGAAAAAATTGGCATTCATCCAGCAGC 609  
Qy 532 ATCGATCCTGTGTAAATTGGAGAACATGGTGATTCCGGTGTGCTGTCTGCTCTTTACC 591  
Db 610 TGCCATGGATGGATTTTGGGGAAACATGGCGACTCAAGTGTGGTGTGGAGTGGTGTG 669  
Qy 592 AACATCGACGGCATGAAGTCTCGGGATTACTCGGAAAAAGCCAAACCATATTTTGATCAG 651  
Db 670 AATGTGCAGGTGTTTCTCTCCAGGAATTGAATCCAGAAATGGGAATGGAATGATAGT 729  
Qy 652 AATGCTTCCATAGAACTTTGAGCAAAACGGAGACGCTGCTTACGATATCATCAAGCGC 711  
Db 730 GAAATTTGGAAGGAAGTGATAAGATGGTGGTTGAAAGTGCCTATGAAGTCATCAAGCTA 789  
Qy 712 AAAGGCTATACCTTATGGAATCGCAGCGGATTACTTCGCATAGTAAAGCGGATTTTA 771  
Db 790 AAAGGATATACCAACTGGGCTATTGGATTAAAGTGTGGTGTGATCTTATTGAATCCATGTTG 849  
Qy 772 GAGGATACAGGATCCACACTTACAGTTTCAACCGTTG 808  
Db 850 AAAATCTATCCAGGATTCTATCCCGTGTCAACAATGG 886

RESULT 6

US-08-869-506-1  
; Sequence 1, Application US/08869506  
; Patent No. 5827710  
; GENERAL INFORMATION:  
; APPLICANT: Uchida, Kohji  
; APPLICANT: Matsukawa, Hirokazu  
; APPLICANT: Matuo, Yushi  
; APPLICANT: Fujita, Teyosi  
; TITLE OF INVENTION: DNA ENCODING THE SUBUNIT OF AVIAN  
; TITLE OF INVENTION: LACTATE DEHYDROGENASE  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 No. 5827710th Glebe Rd. 8th floor  
; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201-4741  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/869,506  
; FILING DATE: 05-JUN-1997  
; CLASSIFICATION: 435





QY 228 TCCTGACTGCGTGGCGGCCCATTTGTTATTTGTCACATGTGGGATTAACCAAAAAAATGG 287  
Db 3294 AGAGGAGGCGAGGCGAGGCAGATATTGTTGTCATCAGCGCGGTGCGAAACAAAGGAAGG 3353  
QY 288 ACAAACAAGGATGATCTTGTCTGCAAAAAATGCCAACATATTATGCTGGAATCATCCCCAA 347  
Db 3354 TGAGCCTCGGACAAAGCTCATTGAACGAACCTTCAGAGTGTGCAAGATCATTTGGTGG 3413  
QY 348 TGTGTCACAAATATGCTCCTGATACCATCCTGCTTATTTGCCACGAATCCTGCTCGATGTTT 407  
Db 3414 CATGCAACCCATTCGACCAGACGCAGTCATCTTGGTGTAGCAATCCAGTCGATATCTT 3473  
QY 408 GACCTATATTAGCTATAAGGCGTCAGGGTTCCACTAAGCAGAGTATCGGCTCAGGTAC 467  
Db 3474 GACACACATTGCAAGACCCCTCTCTGGACTGCTCCAAACCAAGGTCATTTGGCTCCG 3533  
QY 468 AGTTCTGGATCTGCTCGTTTAAATACATCTCTGGAGAGCACTTCAAGATCTCATCGGA 527  
Db 3534 CTACCTTGACAGCACCCGCTCTTCGGCTCCATCTTGGCGATGCTTTGATGTCATCTCA 3593  
QY 528 CAGCATCGATGCTGTGTAATTGGAGAACATGTTGCGGTGTCCTGTCTGTCTCT 587  
Db 3594 ATCGGTCCATGCTTTGCTTGGTGAACATGGGATTCACAGATGCTTGGGAGGC 3653  
QY 588 TACCAACATCGACGGCATGAAGCTCCGGATTAATGCGAAAAAGCCAACCAACATATTTGA 647  
Db 3654 TGCTTCGATTGGTGGCCAGCCGTTTGACAAAGTTTCCCGGAATTCGGAAGC-----TGGAT 3708  
QY 648 TCAGAATGCGTTCATAGAAATCTTTGAGCAACGCGAGAGCGTCTTACGATATCATCAA 707  
Db 3709 AAAACAGCAATTTCAAAGCGATATCAGGTAAGCGA-----TGGAGATCATTCG 3758  
QY 708 GCGCAAAGGCTATACCTTATATGGAATCGCAGCGGATTAATTCGCATAGTAAAGCGAT 767  
Db 3759 TTTGAAAGGAGCGACGTTTATGGAATTTGGTGCCTGTGCAGCGGATTTAGTGACACTAT 3818  
QY 768 TTTAGAGGATACAGGATCCACACTTACAGTTTCAACCGTTGGTGATTTTGGGTTGA 827  
Db 3819 CATGTTGAATAGGAAATCAGTACATCCAGTTTC-----TGTATGTTGAAAGTA 3869  
QY 828 ACAAATGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGCTCATCAAGTGGCTGA 887  
Db 3870 TGGAGCCACTTTTCTATGCTGCTAACTTGGATGGAGAGGTGTTGAACAGATCTATGA 3929  
QY 888 ACTTTCACTCGATGAGAAGGAAATAGAAATGATGGAATAATCAGTAGTCAGATCAATC 947  
Db 3930 AGTACCACCTGACGGAAGAAGAAGCGTTGTTGTAATAATCTGTAGAGGCATTGAAATC 3989  
QY 948 AGT 950  
Db 3990 AGT 3992

RESULT 10  
US-08-270-013B-1  
; Sequence 1, Application US/08270013B  
; Patent No. 5686294  
; GENERAL INFORMATION:  
; APPLICANT: Sogabe et al.  
; TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE  
; TITLE OF INVENTION: DEHYDROGENASE ACTIVITY  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.  
; STREET: Two Prudential Plaza, Suite 4900  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 61601-6780  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/270,013B  
FILING DATE: 01-JUL-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 164701/1993  
FILING DATE: 02-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Green, Robert F.  
REGISTRATION NUMBER: 27555  
REFERENCE/DOCKET NUMBER: 62321  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 616-5600  
TELEFAX: (312) 616-5700  
TELEX: (25)3533  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1912 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Bacillus stearothermophilus  
STRAIN: ATCC12016  
US-08-270-013B-1

Query Match 9.2%; Score 89.2; DB 1; Length 1912;  
Best Local Similarity 51.8%; Pred. No. 1.1e-18;  
Matches 202; Conservative 0; Mismatches 188; Indels 0; Gaps 0;  
QY 223 GATTATCTGACTGCTGCGCGCGGCCCATTTGTTATTTGTCACATGTGGGATTAACCAAAAA 282  
Db 887 GATTACGCTGACACAGCCGATTCGACATCGTCGTCATCAGCAGGCATCGCCCGCAAG 946  
QY 283 AATGACAAACAAGGATGGATCTTCTGCAAAAAATGCCAACATTTATGCTGGAATCATC 342  
Db 947 CCGGCGATGAGCCGCGACGATTTGGTGACGACGAACCAAAAAATTTAAGCAAGTGACG 1006  
QY 343 CCCAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTGCCACGAATCCTGTCGAT 402  
Db 1007 AAGGAAGTCGTCAAATATCTCGCGAACTGCTACATCATCGTCTTGACGAACCCGTCGAT 1066  
QY 403 GTTTTGACCTATATTAGCTATAAGGCGTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCA 462  
Db 1067 GCGATGACGTATACGCTCTTTAAGGAATCCGATTCGGAACCAACCCGTCATCGGCCAG 1126  
QY 463 GGTACAGTTCTGGATACCTGCTGTTTAAATACATCTCGGAGAGCACTTCAAGATCTCA 522  
Db 1127 TCGGCGCTTTGGATACGCGCGGCTTCCGACGTTTCGTCGCGGAGGAGCTGAACATTTG 1186  
QY 523 TCGGACAGCATCGATGCTCTGTAAATTGGAGAACATGTTGATTCGGGTGTCCTGTCG 582  
Db 1187 GTAAAAGATGTCACCTGGGTTGTTTGTAGGCGGCCCATGGCGATGACATGGTGGCGCTCGTC 1246  
QY 583 TCTCTTACCAACATCGACGCGCATGAAGCTC 612  
Db 1247 CGCTACTGTCAGCCGCGGCATTCGCTC 1276

RESULT 11  
US-08-838-418-1  
; Sequence 1, Application US/08838418  
; Patent No. 5744342  
; GENERAL INFORMATION:  
; APPLICANT: Sogabe et al.  
; TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE  
; TITLE OF INVENTION: DEHYDROGENASE ACTIVITY  
; NUMBER OF SEQUENCES: 2



CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leydig, Voit & Mayer, Ltd.  
STREET: Two Prudential Plaza, Suite 4900

Query Match 9.2%; Score 89.2; DB 1; Length 1912;  
Best Local Similarity 51.8%; Pred. No. 1.1e-18;  
Matches 202; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

Qy 583 TCTTTACCAACATCGACGGCATGAAGCTC 612  
 1247 CGCTACTCGTACGCCGGCGGCATTCCGCTC 1276  
 Db

## RESULT 12

```

US-09-134-001C-2696
; Sequence 2696, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGN
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064, 964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055, 779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2696
; LENGTH: 978
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2696

```

Query Match 8.3%; Score 81; DB 4; Length 978;  
Best Local Similarity 51.5%; Pred. NO. 3.3e-16;  
Matches 186; Conservative 0; Mismatches 175; Indels 0; Gaps 0

## RESULT 13

US-09-328-111-455  
; Sequence 455, Application US/09328111  
; Patent No. 6262333  
; GENERAL INFORMATION:  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Steinmann, Kathleen E.  
; APPLICANT: Astle, Jon H.  
; APPLICANT: Burgess, Christopher C.  
; APPLICANT: Bushnell, Steven E.  
; APPLICANT: Carroll III, Eddie  
; APPLICANT: Catino, Theodore J.  
; APPLICANT: Derti, Adnan



Job time : 65 secs

APPLICANT: Uehara, Hirotugu  
APPLICANT: Eda, Soji  
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/920,827  
FILING DATE: 29-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,577  
FILING DATE: 27-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 19036/32420  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5024 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
ORIGINAL SOURCE:  
ORGANISM: Staphylococcus epidermidis  
STRAIN: Clinical Isolate SE-32

US-08-920-827-7

Query Match 7.6%; Score 74.2; DB 1; Length 5024;  
Best Local Similarity 51.7%; Pred. No. 1.2e-13;  
Matches 169; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

QY	279	AAAAAATGGACAAACAAGGATGGATCTTGCTGCAAAAAAATGCCAACATTATGCTGGAAT	338
Db	306	AAATCAGGAATGACAAGGAGAAGATTAGTTCAAACCTAATGAACAATAGTACGAGAAC	365
QY	339	CATCCCCAATGTTGCCAAATATGCTCCTGTATACCATCCTGCTTATTGCCACGAATCCTGT	398
Db	366	TGCATTACAAATTGCAACGTATGCACCTCATTCATAATTATTGTATTGACTAATCCGGT	425
QY	399	CGATGTTTTGACCTATATTAGCTATAAGGCGTCAGGGTTTCCACTAAGCAGAGTTATCGG	458
Db	426	TGATGTTATGACATATACTGCAATTTAAAGCATCAGGTTTTCTTAAAGAACGTATTATTGG	485
QY	459	CTCAGGTACAGTTCTGGATACCTGCTCGTTTTTAAATACATCCTCGGAGAGCACTTCAAGAT	518
Db	486	TCAATCTGGAATTTTAGACGCTGCAAGATATCGAACTTTTATTGCTCAAGAACTTAACGT	545
QY	519	CTCATCGGACAGATCGATGCCTGTGTAATTGGAGAAACATGGTGATTGGGGTGTGCCTGT	578
Db	546	GTCTGTCAAAGATGTAAATGGGTTTGTGTTTAGGTGGACATGGTGATACCATGTTACCTTT	605
QY	579	CTGGTCTCTTACCAACATCGCGGCAT	605
Db	606	GATTATAACACACACATTAATGGAT	632

Search completed: July 31, 2003, 11:02:32

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 09:47:41 ; Search time 260 Seconds  
(without alignments)  
7712.449 Million cell updates/sec

Title: US-09-992-430B-21  
Perfect score: 972  
Sequence: 1 atgttccaagatacaaaagtc.....ttgagcatctggagatcaat 972

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
  
Searched: 1439767 seqs, 1031500376 residues  
Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq1.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq3.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	186.2	19.2	28690	11	US-09-070-927A-138
2	183.6	18.9	1255	10	US-09-971-361-2
3	173	17.8	13154	11	US-09-070-927A-338
4	172.6	17.8	3264	8	US-08-781-986A-424
5	159.8	16.4	14333	8	US-08-781-986A-57
6	154.2	15.9	942	11	US-09-738-626-3210
7	154.2	15.9	3309400	11	US-09-738-626-1
8	151	15.5	747	11	US-09-974-300-73
9	136	14.0	1146	15	US-10-274-266-1
10	136	14.0	1759	11	US-09-731-872-14
11	136	14.0	3144	15	US-10-274-266-3
12	123.6	12.7	1062	11	US-09-938-842A-1938
13	121.2	12.5	4913	10	US-09-971-361-1
14	120.8	12.4	1755	11	US-09-731-872-15
15	109.8	11.3	1172	11	US-09-974-298-74
16	109.8	11.3	1272	15	US-10-177-293-261

17	109.8	11.3	2719	15	US-10-198-846-11544	Sequence 11544, A
18	109.8	11.3	3397	15	US-10-198-846-11555	Sequence 11555, A
19	108.2	11.1	2719	15	US-10-198-846-11544	Sequence 11544, A
20	106.4	10.9	491	11	US-09-960-352-9428	Sequence 9428, Ap
21	102.2	10.5	734	10	US-09-910-943-711	Sequence 711, App
22	102	10.5	1362	11	US-09-974-298-73	Sequence 73, Appl
23	102	10.5	1362	15	US-10-044-090-556	Sequence 556, App
24	102	10.5	1362	15	US-10-071-766-89	Sequence 89, Appl
25	100.6	10.3	496	11	US-09-998-598-1409	Sequence 1409, Ap
26	100.6	10.3	519	11	US-09-998-598-1410	Sequence 1410, Ap
27	85	8.7	822	11	US-09-974-300-56	Sequence 56, Appl
28	83.2	8.6	839	15	US-10-198-846-2972	Sequence 2972, Ap
29	83.2	8.6	883	15	US-10-198-846-8917	Sequence 8917, Ap
30	79.8	8.2	566	11	US-09-879-536-455	Sequence 455, App
31	76	7.8	891	15	US-10-198-846-2988	Sequence 2988, Ap
32	74.2	7.6	379	11	US-09-960-352-14665	Sequence 14665, A
33	72.6	7.5	974	15	US-10-198-846-9276	Sequence 9276, Ap
34	70.2	7.2	520	11	US-09-998-598-1812	Sequence 1812, Ap
35	70.2	7.2	893	15	US-10-198-846-8905	Sequence 8905, Ap
36	69.2	7.1	530	11	US-09-833-381-773	Sequence 773, App
37	66.4	6.8	394	11	US-09-960-352-7674	Sequence 7674, Ap
38	66.4	6.8	436	11	US-09-736-457-1076	Sequence 1076, Ap
39	66.4	6.8	436	11	US-09-902-941-1076	Sequence 1076, Ap
40	66.4	6.8	436	11	US-09-849-626-1076	Sequence 1076, Ap
41	66.4	6.8	436	15	US-10-017-754-1076	Sequence 1076, Ap
42	63	6.5	422	11	US-09-920-300A-1013	Sequence 1013, Ap
43	63	6.5	422	15	US-10-033-528-1013	Sequence 1013, Ap
44	62.2	6.4	519	10	US-09-604-287A-160	Sequence 160, App
45	62.2	6.4	519	11	US-09-339-338-160	Sequence 160, App

ALIGNMENTS

RESULT 1

US-09-070-927A-138

; Sequence 138, Application US/09070927A  
; Patent No. US20020120116A1

GENERAL INFORMATION:

APPLICANT: Charles A. Kunsch  
Patrick J. Dillon  
Steven Barash

TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 982

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/070,927A

FILING DATE: 04-May-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/046,655

FILING DATE: 1997-05-16

APPLICATION NUMBER: 60/044,031

FILING DATE: 1997-05-06

APPLICATION NUMBER: 60/066,009

FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB369

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504



```

; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28690 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-09-070-927A-138

Query Match      19.2%; Score 186.2; DB 11; Length 28690;
Best Local Similarity 52.5%; Pred. No. 5.1e-45;
Matches 481; Conservative 1; Mismatches 420; Indels 14; Gaps 3;

QY 51 AAAAGTTGTGGTAGTGGAGTGGGAAGTGTGGTCTGCCACAGCGTATACGTTGCTTCT 110
Db 16610 AAAAGTCGCAATTATGGTACTGGTTTGTGGCACAAGTATTGCCCTATTCATGATCAA 16669

QY 111 CAGCGGCATCGTTTCCGAGATTGTCCTTATCGAGCTGAACAAAGACAAAGCAGAGGGTGA 170
Db 16670 CCAAGGGATTGCGAATGAATTAATCTTAGTTGATATTGACAAAGCCAAATCTGAAGCGCA 16729

QY 171 AAGCATGGACTTAAACACG-----CAGCACCTTCAAATACAAGGTCTCGAGCGGGT 222
Db 16730 NCAATTGACTTATTAGATGGTGTGKCYTGGGKCMAGAAATGTAAACGTCTGGGCTGGC 16789

QY 223 GATTATCCTGACTGCGCTGGCGGGCCATTGTTATTGTTCACATGTGGGATTAACCAAAA 282
Db 16790 GACTATCAAGACTGCCAAGATGCCGATATCGTCTGTATTACAGCTGGCGCTAATCAAAA 16849

QY 283 AATGGACAAACAAGGATGGATCTTGCTGCAAAAAAATGCCAACATTTATGCTGGAATCATC 342
Db 16850 CCTGGGCAAGCTCGTCTAGATTGGTTTCAATCAATGCAGAAATTTATGAAAAAATGTT 16909

QY 343 CCCAATGTTGCCAAATATGCTCCTGTATACCATCCTGCTTATTGCCACGAATCCTGTCTGAT 402
Db 16910 AACAAATATCATGAATCTGGTTTGTATGGAATTTTAGTGATTGGCTCAAAATCCTGTCTGAT 16969

QY 403 GTTTTGGACCTATATTAGCTATAAGGCGTCAAGGTTTCCACTAAGCAGAGTATCGGCTCA 462
Db 16970 GTACTGACTTATGIGGCTGGCAAGCTTCTGTTTACCTGTTTCAAGAGTAATTGGAAC 17029

QY 463 GGTACAGTCTCGGATACCTCGTCTTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCA 522
Db 17030 GGTACAACTTTGGACACAACTCGTTTCCGCAAGAAGTGTCTCAACGTTTAGCGATTGAT 17089

QY 523 TCGGACAGCATCGATGCTGTGTAAATTGGAGAACATGGTGTATTCGGGTGTCCTGTCTGG 582
Db 17090 CCACGCAATGTTACGCTATATTATTGGCGAACACGGGATTTCTGAAGTGGCTGTCTGG 17149

QY 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGGAAAAAGCCAACACATA 642
Db 17150 TCTCATACCATGATTGGTACCAACCTATTTTAGAAAATTGTGGATACGACAGAGCGCTTA 17209

QY 643 TTTGATCAGAAATCGGTTCCATAGAAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATATC 702
Db 17210 ACTAGTGACGAT---TTACCAATCATTTCTGATAAAGTGAAAAATACAGCTTATGAAATT 17266

QY 703 ATCAAGCGCAAGGCTATACCTTATATGGAATCGCAGCGGGATTACTTCGATAGTAAAG 762
Db 17267 ATCGATCGCAACACAGCGACCTATTATGGAATTTGGTATGAGTACTGACGCAATGTTAAA 17326

QY 763 GCGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAAC---CGTTGGTGAATTTT 819
Db 17327 GCCATTTTAAATAATGAACAAGCTATTTTACCTGCTCTCAGCTTACTTAGATGGCAATAT 17386

QY 820 GGGGTTGAACAAATTGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGCTCATCAA 879
Db 17387 GGGCAACAAGATGTATTACAGGGATTCTCTGCAATCGTTGGCAATCAAGGTGTGACTGAC 17446

QY 880 GTGGCTGAACCTTTCACCTGATGAGAAGGAATAGAAATTGATGGAAAAATCAGCTAGTCAG 939
Db 17447 ATTATTGAATTGAATCTGAATGCCGCTGAAAAAGAACTCTTCCAAAAATCAGTGACACAA 17506
```

```

QY 940 ATCAATCAGTGATTG 955
Db 17507 TTAAACAAGTGATGG 17522

RESULT 2
US-09-971-361-2
; Sequence 2, Application US/09971361
; Patent No. US20020081677A1
; GENERAL INFORMATION:
; APPLICANT: Javed, Muhammad
; APPLICANT: Cusdin, Fiona
; APPLICANT: Milner, Paul
; APPLICANT: Green, Edward
; TITLE OF INVENTION: Ethanol Production
; FILE REFERENCE: 000487.00010
; CURRENT APPLICATION NUMBER: US/09/971,361
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/247,017
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: UK 0024554.8
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: DNA
; ORGANISM: Bacillus LN
US-09-971-361-2
```

```

Query Match      18.9%; Score 183.6; DB 10; Length 1255;
Best Local Similarity 53.2%; Pred. No. 6e-45;
Matches 465; Conservative 0; Mismatches 394; Indels 15; Gaps 3;

QY 54 AGTTGTGGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTCAG 113
Db 322 AGTAGCACTTATAGGAACGGGGTTCGTTGGGCCAGCTATGCATTTGCCCTTATGAACCA 381

QY 114 CGGCATCGTTTCCGAGATTGTCCTTATCGACGTGAACAAAGACAAAGCAGAGGGTGAAG 173
Db 382 AGGAATAGCAGATGAGTTAGTATTGATGTAAATAAGATAAGGCAGAGGGCGATGT 441

QY 174 CATGACTTAAACCCAGCAGCACTTCAAAATACAAGGTC-----TCGAGCGGGTGA 224
Db 442 GATGGATTTAAATCACGGAAAAGTATTCGCGCCGAAGCCGATGAATATTGTTGGAGA 501

QY 225 TTATCCTGACTGCGTGGCGGGCCATTGTTATTGTACATGTGGGATTAAACCAAAAAA 284
Db 502 TTATCAAGATTGCCAAGACGCGGATTTGGTGGTGTATTTGTGACGGGGCTAACCAAAAGCC 561

QY 285 TGGACAAACAAGGATGGATCTTGTGTCAAAAAATGCCAACATTTATGCTGGAATCATCCC 344
Db 562 GGGAGAAACAAGACTGGATCTTTGTGACAAAAATATTAATATCTTCAAAACGATTGTGCA 621

QY 345 CAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTGCCACGAATCCTGTGATGT 404
Db 622 TTCTGTGATGAATCCGGATTGTGATGGCGTTTCTTGTGGCAACGAACCCAGTGGATAT 681

QY 405 TTTGACCTATATTAGCTATAAGCGGTCAAGGTTTCCACTAAGCAGAGTTATCGGCTCAGG 464
Db 682 TTTAACGTATGCTACTTGGAAATTTAGCGGGTTACCGAAAGAGCGGGTAAATCGGCTCAGG 741

QY 465 TACAGTTCTGGATACCTGCTGTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCATC 524
Db 742 AACGATTTTGATACAGCAAGATTCGCTTCTTGTCTAAGTGAATATTTTCAAGTGGCTCC 801

QY 525 GGACAGCATCGATGCTGTGTAAATTGGAGAAACATGGTATTGGGTGTGCTGTGGTC 584
Db 802 GACCAATGTACATGCGTATATTATTGGCGAGCATGGGGATACAGAGTGCCTGTTTGGAG 861

QY 585 TCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCAAACCATATT 644
```

Db 862 CCATCGGGAATTGGAAGCATTCCAGTTGAGCAAAATATTGATGCAA---AACGATAACTA 918  
QY 645 TGATCAGAAATGCGTTCCATAGAATCTTTGAGCAAAACGGAGACGCTGCTTACGATATCAT 704  
Db 919 TAGAAAAGAGGATTAGACAATATCTTTGTTAAATGTTTCGTATGCGGCATATCAAAATCAT 978  
QY 705 CAAGCGGAAAGGCTATACCTTATATGGAATCGCAGCGGGATTACTTTCGCATAGTAAAGGC 764  
Db 979 TGAGAAAAAGGGCAACGATTACGGCATGCAATGGGATTAGTCCGTATCACTCGTGC 1038  
QY 765 GATTTTAGAGGATACAGGATCCACACTTACAGTTTCAAC---CGTTGGTGATTATTTGG 821  
Db 1039 TATTTTGCACAATGAAAATGCCATCTTAACCGTTTCTGCTCATTTGGACGGCCAATATGG 1098  
QY 822 GGTTGAACAAATGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGGCTCATCAAGT 881  
Db 1099 CGAACGAAATGTTTATATTGGCGTGCCTGCCATTATCAACCGAAACGGTATTTCGTGAAGT 1158  
QY 882 GGCTGAACCTTTCACCTCGATGAGAAGGAAATAGAA 915  
Db 1159 GATGGAATTGACGCTAAATGAAACAGAACAA 1192

RESULT 3  
US-09-070-927A-338/c  
; Sequence 338, Application US/09070927A  
; Patent No. US20020120116A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Kunsch  
; Patrick J. Dillon  
; Steven Barash  
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 982  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/070,927A  
; FILING DATE: 04-May-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/046,655  
; FILING DATE: 1997-05-16  
; APPLICATION NUMBER: 60/044,031  
; FILING DATE: 1997-05-06  
; APPLICATION NUMBER: 60/066,009  
; FILING DATE: 1997-11-14  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kenley K. Hoover  
; REGISTRATION NUMBER: 40,302  
; REFERENCE/DOCKET NUMBER: PB369  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 338:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13154 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 338:

Query Match 17.8%; Score 173; DB 11; Length 13154;

Best Local Similarity 51.7%; Pred. No. 3.4e-41;  
Matches 473; Conservative 0; Mismatches 430; Indels 12; Gaps 3;  
QY 51 AAAAGTTGTGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCT 110  
Db 6674 AAAAGTAAATTTAGTCGGGGAGCGTGCCGTAGGTTCTAGCTATGCTTTCCTTTAGTAAC 6615  
QY 111 CAGCGGCATCGTTTCCGAGATTGTCTTATCGACGTGAACAAAGACAAAGCAGAGGGTGA 170  
Db 6614 TCAGAATATTGCTCAAGAAAGTTGGGATTATTGATATTAAATGTACCAAAAACCTGAAGGAGA 6555  
QY 171 AAGCATGGACTTAAACCACGCGCAC-----CTTCAAATACAAGGTCTCGAGCGGGTGA 224  
Db 6554 CGCGTTGGACTTATCTCACGCATTAGCATTTACTTCTCTAAAAAAATCTATGCTGTCTAC 6495  
QY 225 TTATCCTGACTGCGTGGCGGGCCATTTGTTATTGTCAATGTGGGATTAAACCAAAAAA 284  
Db 6494 TTATGACGATTGCCATGATGCAGACTTAGTTGTCTTAACAGCTGGTGGCCTCAAAAAACC 6435  
QY 285 TGGACAAAAAAGGATGGATCTTGTCTGCAAAAAAATGCCAACATTATGCTGGAATCATCCC 344  
Db 6434 AGCGAAAACTCGTTTAGACTTAGTTTATATAAACTTGAAAAATTAATAAAGAAATCGTTAC 6375  
QY 345 CAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTGCGACGAATCCTGTCGATGT 404  
Db 6374 AACAAATTGTTGATTCTCGTTTCAACGGTATCTTCTTAGTTGCCGCAAAACCCAGTTGATAT 6315  
QY 405 TTTGACCTATATTAGCTATAAGCGCTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCAGG 464  
Db 6314 TTTGACTTATTCAACTTGGAAATCTCTGGCTTCCCGAAAGAGAGTAATCGGTTTCAGG 6255  
QY 465 TACAGTTCTGGATACGTCGTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCATC 524  
Db 6254 AACTTCACTAGATTCTGCTCGTTCCGTCAAGCAATGCGGAATTAGTTGACGTTGATGC 6195  
QY 525 GGACAGCATCGATGCTGTGTAATTGGAGAACATGGTGATTCCGGTGTCCCTGTCTGGTC 584  
Db 6194 ACGAAATGTCCATGCCTACATCTTAGGGGAACACGGAGATACAGAATTTCCAGTTTGGTC 6135  
QY 585 TCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCAAACCATATT 644  
Db 6134 ACATCGGAATGTCGCTGGCTTACAAAT---TTACGAATGGGTGAAAAATAATCTCGACGT 6078  
QY 645 TGATCAGAAATGCGTTCCATAGAATCTTTGAGCAAAACGGCAGACGCTGCTTACGATATCAT 704  
Db 6077 CGATGAAGAAGCAATGGTTAATTATTCTTCAACGTACGCGACGCTGCTTACACAATCAT 6018  
QY 705 CAAGCGCAAAAGGCTATACCTTCAATGGAATCGCAGCGGATTACTTCGCATAGTAAAGGC 764  
Db 6017 CGAGAAAAAAGGAGCTACTTTCTATGGAATCGCGGTTGCACTAGCGGTATCACTAAAGC 5958  
QY 765 GATTTTAGAGGATACAGGATCCACACTTACAGTTTC---AACCGTTGGTGATTATTTGG 821  
Db 5957 TATCCTAAACGATGAAAACTCTGTGTACCATTATCTGTTTATTAGAGGTGAATATGG 5898  
QY 822 GGTGAACAAATTTGCTATAAGCGTCCCTACCACCAACTCAATAAAAGTGGGGCTCATCAAGT 881  
Db 5897 TCAAAACGATATTTATATCGGTGCACCCAGCGATCATCAACCGCAAGGAGTTAAACAAGT 5838  
QY 882 GGCTGAACCTTTCACTCGATGAGAAGGAAATAGAAATTGATGGAAAAATCAGCTAGTCAGAT 941  
Db 5837 CATTGAAATTCATTAAACAGATGCTGAAACAAGAAAAAATGGAAGCTTCTGCTTCTGCATT 5778  
QY 942 CAAATCAGTGATTGA 956  
Db 5777 AAAAGAAGTTATTGA 5763

RESULT 4  
US-08-781-986A-424  
; Sequence 424, Application US/08781986A  
; Publication No. US20030054436A1  
; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 424:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-424

Query Match 17.8%; Score 172.6; DB 8; Length 3264;
Best Local Similarity 52.2%; Pred. No. 2.1e-41;
Matches 469; Conservative 0; Mismatches 409; Indels 21; Gaps 3;

QY 51 AAAAGTTGTGTAGTGGGAGTGGGAAGTTGGGTCTGCCACAGCGTATACGTTGCTTCT 110
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
412 AAAGGTTGTATTAAATCGGAGATGGATCTGTAGGATCAAGCTATGCCTTTGCAATGGTTAC 471

QY 111 CAGCGGCATCGTTTCCGAGATTGTCTTATCCAGCTGAACAAAGACAAAGCAGAGGTGA 170
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
472 GCAAGGTGTGCTGATGAATTGTAAATTATTGACATTGCAAAAGACAAAGTAAAAGCAGA 531

QY 171 AAGCATGGACTTAAACCACGCAGCA-----CCTTCAAATACAAAGTCTCGAGCGGG 221
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
532 TGTTCAAGATTTAAACCATGGTACAGTCCACAGTCCTTCCACCAGTTGATGTGAAGCAGG 591

QY 222 TGATTATCCTGACTGCGCTGGCGCGGCCATTGTTATTGTCACATGTGGGATTAAACCAAA 281
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
592 TGAATACGAAGACTGTAAAGATGCAGATTAGTTAGTTATTACAGCTGGTGCACTCAAAA 651

QY 282 AAATGGACAAACAAAGGATGGATCTTGCTGCAAAAATGCCAACATTATGCTGGAAATCAT 341
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
652 GCCAGGTGAACACCGTTTACAATTAGTTGAAAAAAATACTAAGATTATGAAGAGCATCGT 711

QY 342 CCCCAATGTTGCCAAATATGCTCCTGATACCATCCTGTTATTGCCAGGAATCCTGTGGA 401
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
712 TAAGAGTGTATTGGATAGTGGCTTTTGATGGATATTCTTAATCGCGCAAAACCTGTAGA 771

QY 402 TGTTTGGACCTATATTAGCTATAAGCGGTTCAGGTTTCCACTAAGCAGAGTTATCGGCTC 461
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
772 CATTTTAAACAAGATTGTAAAGAATATATACTGGATTACCAGCAGAGCGTGTATCGGTTTC 831

QY 462 AGGTACAGTTCTGGATACTGCTCGTTTTTAATATACATCCTCGGAGAGGCACTTCAAGATCTC 521
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
832 AGGTACTGTATTGGACAGTGCACGTTTACATAATTAAATTAGCCCAAGAACTTGGTGTTC 891

QY 522 ATCGGACAGCATCGATGCCTGTGTAATTGGAGAAACATGGTGAATTCGGGTGTGCTGTCTG 581
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
892 ACCTTCAAGTGTGACGCTAGTATTATTGGCGAGCATGGTGATACTGAACCTTGACAGTTTG 951

QY 582 GTCTCTTACCAACATCGACGCGCATGAAGCTCCGGGATTACTCGGAAAAAGCCAACCAT 641
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
952 GTCACAAGCAAATGTAGCAGGTATTTTCAGTATATGACACATAAAAGAA-----CAAC 1005

QY 642 ATTTGATCAGAATGCGTTTCCATAGAAATCTTTGAGCAACCGGAGACGCTGTACGATAT 701
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1006 TGGTAGCGAAGCTAAAGCGGAAGAAATTTATGTGAATACACGTGACGCTGCTTATGAAAT 1065

QY 702 CATCAAGCGCAAAGGCTATATCTTATGGAATCGCAGCGGATTAATTCGCATAGTAA 761
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1066 TATCCAAGCTAAAGGGTCAACATACTATGTTGTTAGTATGCGCATTTTCAAA 1125

QY 762 GCGGATTTA-----GAGGATACAGGATCCACACTTACAGTTTCAACCGTTGGTGATTA 815
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1126 AGCCATTTTAAATAATGAATAAATGTCTTAAATGTTTCTATACAATTAGATGGTCAATA 1185

QY 816 TTTTGGGGTTGAACAAATTCGTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGCTCA 875
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1186 TGGTGGTCACAAAGCGGTTTACCTAGGTGTACCAACATTAGTTAACCAACATGGCGCAGT 1245

QY 876 TCAAGTGGCTGAACTTTTCACTCGATGAGAGGAAGAAATAGAAATGATGGAAAAATCAGCTA 934
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1246 TAAATTTATGAAATGCCATTAAAGTGCCGAAGAACAAAGCGTTGTTGATAAATCTGTTA 1304

RESULT 5
US-08-781-986A-57/c
; Sequence 57, Application US/08781986A
; Publication No. US2003005436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14333 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-57

Query Match 16.4%; Score 159.8; DB 8; Length 14333;
Best Local Similarity 51.2%; Pred. No. 3.6e-37;



	Matches	481;	Conservative	1;	Mismatches	439;	Indels	19;	Gaps	4;
QY	28	GTAAGAACTGATG	CCAAAACAGTAAAAAGTTGTGGTAGTGGAGTGGGAAGTGTGGGTCT	87						
Db	13342	GTAATGAACAAATTTAAAGGGAACAAAGTTGTATTAAATAGGTAATGGTGCAGTAGGTTCA	13283							
QY	88	GCCACAGCGTATACGTTGCTTCTCAGCGGCATCGTTTCCGAGATTGTCTTATCGACGTG	147							
Db	13282	AGCTACGCATTTTCATTAGTGAACCAAGCATTTGTTGATGAATTAGTCATCATTTGATTTA	13223							
QY	148	AACAAAGACAAAGCAGAGGGTGAAGCATGGACTTAAACACCGCAGCACCTTCAAAATACA	207							
Db	13222	GACACTGA AAAAGTTTCGAGGAGATGTTATGGATTTAAACATGCCACACCATATTTCTCCA	13163							
QY	208	AGGTCTCG-----AGCGGTGATTATCCTGACTGCGCTGGCGCGGCCATTGTTATT	258							
Db	13162	ACAACAGTTCGTGTGAAAGCTGGCGAATACAGTGATTGTCTATGATCGGATCTAGTTGTCT	13103							
QY	259	GTCACATGTGGGATTAAACCAAAAAATGGAACAAAGGATGGATCTTGTGCAAAAAAAT	318							
Db	13102	ATCTGTGCTGGTGTGCACAAAAACCTGGAGAAACACGTTTAGATTAGTATCTAAAAAC	13043							
QY	319	GCCAAATATATGCTGAAATCATCCCAATGTTGCCAAATATGCTCCTGATACCATCTG	378							
Db	13042	TTGAAAAATATTCAAATCAATTGTTGGTGAAGTAATGGCATCAAAATTTGATGGTATTTTC	12983							
QY	379	CTTATTGCCACGAATCCTGTCTGATGTTTGGACCTATATTTAGCTATAAGGCGTCAGGGTTT	438							
Db	12982	TTGGTAGCTACAAATCCTGTTGATATTTAGCGTATGCAACATGGAAATTTCTCTGGTTTA	12923							
QY	439	CCACTAAGCAGAGTTATCGGCTCAGGTACAGTTCTGGATACTGCTCGTTTTTAAATACATC	498							
Db	12922	CCTAAAGAACGTGTTATAGTTCTGGTACAAATTTTAGACTCTGCACGCTTTAGATTATTG	12863							
QY	499	CTCGGAGAGCATTCAAGATCTCATCGGACAGCATCGATGCTGTGTAATTGGAGAACAT	558							
Db	12862	TTAAGCGARC-GTNCGATGTTGCGCCACGTAGCGTCAAAATTTATTTGGTGAACAT	12804							
QY	559	GGTGATTCCGGTGTGCCCTGTCTGCTCTTACCAACATCGACGGCATGAAGTCCGGGAT	618							
Db	12803	GGTGACACTGAATTACCAAGTATGGTCACACGCTAATATTGCGGGTCAACCTTTGAAGACA	12744							
QY	619	TACTGCGAAAAAGCCAAACCATATTTTGATCAGAAATCGTTCATAGAATCTTTTGACAA	678							
Db	12743	TTACTTTGAACAACGTCCTGAGGGGCAAGCGCAAATTG-----AACAAATTTTGTTCAA	12690							
QY	679	ACGCGAGACGCTGCTTACGATATCATCAAGCGCAAGGCTATACTTCATATGGAATCGCA	738							
Db	12689	ACACGTATGCAGCATATGACATTATTCAAGCTAAAGGTGCCACTTATTATGGTGTGCA	12630							
QY	739	GCGGATTACTTCGCATAGTAAAGCGCATTTTAGAGGATACAGGATCCACACTTACAGTT	798							
Db	12629	ATGGGATTAGCTAGAAATTACTGAAGCGATTTTCAGAAATGAAGATGCCGATTGACTGTA	12570							
QY	799	TCAAC---CGTTGGTGATTATTTTGGGGTTGAACAAATTCGTATAAGCGTCCCTACCAAA	855							
Db	12569	TCAGCATTTATTAGAAGGCGAATATAGGAAGAAGATGTTTATATTGGTGTCCAGCAGTC	12510							
QY	856	CTCAATAAAAAGTGGGGCTCATCAAGTGGCTGAACTTTTCACTCGATGAGAAGGAATAGAA	915							
Db	12509	ATCAATAGAACCGGTATTCGCAACGTCGTAGAAATCCCATTAACGACGAAGAACAAAGC	12450							
QY	916	TTGATGGAAAAATCAGCTAGTCAGATCAAAATCAGTGATTG	955							
Db	12449	AAGTTCGCACATTCAGCTAAAAACATTAAAAAGATATTATGG	12410							

; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAAKO  
 ; APPLICANT: SENOH, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 3210  
 ; LENGTH: 942  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-3210

Query Match	15.9%;	Score 154.2;	DB 11;	Length 942;
Best Local Similarity	50.8%;	Pred. No. 4.2e-36;		
Matches 456;	Conservative 0;	Mismatches 423;	Indels 18;	Gaps 3
QY	52	AAAGTTGTGGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC	111	
DB	22	AAGATTGTCTCATTTGGCGCAGGAGATGTTGGAGTTGCATACGCATACGCACTGATCAAC	81	
QY	112	AGCGGCATCGTTTCCGAGATTGTCTTATCGACGTGAACAAAGACAAAGCAGAGGGTGAA	171	
DB	82	CAGGGCATGGCAGATCACCTTTCGATCATCGACATCGATGAAAGAACTCGAAGGCAAC	141	
QY	172	AGCATGGACTTAAACCAACGCA-----GCACCTTCAAATACAAGTCTCGAGCGGGT	222	
DB	142	GTCAATGGACTTAAACCAATGGTGTGTGGGCCGATTCCCGCACCCGCGTACCAAGGCG	201	
QY	223	GATTATCCTGACTGCGTGGCGCGGCCATTGTTATTGTACATGTGGATTAAACCAAAAA	282	
DB	202	ACCTACGTGACTGCGAAGACGACGCCATGGTTGTCAATTTGTGCCGCGCAGCCCCAAAAG	261	
QY	283	AATGGACAAACAAGGATGGATCTTGTCTGCAAAAAATGCCAACATTATGCTGGAAATCATC	342	
DB	262	CCAGCGGAGACCCGCCCTCCAGTGGTGACAAAAACGTCAAGATTATGAAATCCATCGTC	321	
QY	343	CCCAATGTTGCCAAATATGCTCTGTATACCATCCTGCTTATTGCCACGAATCCTGTCTGAT	402	
DB	322	GGCGATGTCATGGACAGCGGATTCGACGGCATCTTCTCGTGGCGTCCAACCCAGTGGAT	381	
QY	403	GTTTTTGACCTATATTAGCTATAAGGCGTCAAGGTTTCCACTAAGCAGAGTTATCGGCTCA	462	
DB	382	ATCCTGACCTACGCAGTGTGGAATTTCTCCGGCTTGGAAATGGAACCGCGTGATCGGCTCC	441	
QY	463	GGTACAGTTCTGGATACTGCTCGTTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCA	522	
DB	442	GGAACCTGCTTGACTCCGCTCGATTCCGTACATGCTGGCGGAATCTTACGAAGTGGCA	501	
QY	523	TCGGACAGCATCGATGCTGTGTAATTGGAGAACATGGTGATTCCGGGTGTGCCTGTCTGG	582	
DB	502	CCAAGCTCCGTCCACGCCTACATCATCGGCGAACACGGCGACACTGAACTTCCAGTCCCTG	561	
QY	583	TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCAACACCATATA	642	
DB	562	TCCTCCGCGACCATCGCAGGCGTATCGCTTAGCCGAATGCTGGACAAAGACCCAGAGCT-	620	
QY	643	TTTGATCAGAAATGCGTTCCATAGAATCTTTTGAGCAACCGGAGACGCTGCTTACGATATC	702	



Db 621 -----TGAGGGCCGCTAGAGAAAATTTTCGAAGACACCCGCGACGCTGCCTATCACATT 675

QY 703 ATCAAGCGCAAAGGCTATACCTTCATATGGAATCGCAGCGGGATTACTTCGCATAGTAAAG 762

Db 676 ATCGACGCCAAGGCTCCACTTCTACGGCATCGGCATGGGTCTTGTCTGCATCACCCGC 735

QY 763 GCGATTTTAGAGGATACAGGATCCACACTTACAGT---TTCAACCGTTGGTGATTATTT 819

Db 736 GCAATCCTGCAGAACCAAGACGTTGCAGTCCAGTCTCTGCACTGCTCCACGGTGAATAC 795

QY 820 GGGGTTGAACAAATTGCTATAGCGTCCCTACCAAACTCAATAAAAGTGGGGCTCATCAA 879

Db 796 GGTGAGGAAGACATCTACATCGGCACCCAGCTGTGTGTAACCGCCGAGGCATCCGCCGC 855

QY 880 GTGGCTGAACCTTTCACTCGATGAGAAGGAATAGAAATTGATGGAATAATCAGCTAGT 936

Db 856 GTTGTGCAACTAGAAATCACCGACACGAGATGGAACGCTTCAAGCATTTCCGCAAT 912

RESULT 7

US-09-738-626-1/c

; Sequence 1, Application US/09738626

; Publication No. US20020197605A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI

; APPLICANT: MIZOGUCHI, HIROSHI

; APPLICANT: ANDO, SEIKO

; APPLICANT: HAYASHI, MIKIRO

; APPLICANT: OCHIAI, KEIKO

; APPLICANT: YOKOI, HARUHIKO

; APPLICANT: TATEISHI, NAKO

; APPLICANT: SENOH, AKIHIRO

; APPLICANT: IKEDA, MASATO

; APPLICANT: OZAKI, AKIO

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-125

; CURRENT APPLICATION NUMBER: US/09/738,626

; CURRENT FILING DATE: 2000-12-18

; PRIOR FILING DATE: 2000-12-16

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: JP 99/377484

; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: JP 00/159162

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: JP 00/280988

; NUMBER OF SEQ ID NOS: 7059

; SOFTWARE: PatentIn ver. 3.0

; SEQ ID NO 1

; LENGTH: 3309400

; TYPE: DNA

; ORGANISM: Corynebacterium glutamicum

US-09-738-626-1

Query Match 15.9%; Score 154.2; DB 11; Length 3309400;

Best Local Similarity 50.8%; Pred. No. 3.2e-34;

Matches 456; Conservative 0; Mismatches 423; Indels 18; Gaps 3;

QY 52 AAAGTTGTGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC 111

Db 3113369 AAGATTGTCTCAITTTGGCGCAGGAGATGTTGGAGTTGCATACGCATACGCACTGATCAAC 3113310

QY 112 AGCGGCATCGTTTCCGAGATTGCTCTTATCGAGCTGAACAAAGACAAAGCAGAGGGTGAA 171

Db 3113309 CAGGGCATGGCAGATCACCTTGGATCATCGACATCGATGAAAGAAACTCGAAGGCAAC 3113250

QY 172 AGCATGGACTTAACCAACCGCA-----GCACCTTCAAATACAAGGTCTCGAGCGGT 222

Db 3113249 GTCATGGACTTAACCACTGTTGTGTGGGCGGATTCCCGCACCCGCGTCAACCAAGGGC 3113190

QY 223 GATTATCTGACTGCGTGGCGCGCCATTGTTATTGTCACATGTGGGATTAACCAAAAA 282

Db 3113189 ACCTACGCTGACTGCGAAGACGACGCCATGTTGTTCATTGTCGCGCGCAGCCCAAAAG 3113130

QY 283 AATGGACAAACAAGGATGATCTTGTCTGCAAAAAAATGCCAACATTATGCTGGAATCATC 342

Db 3113129 CCAGCGAGACCCGCTCCAGTGGTGACAAAAACGTCAAGATTATGAAATCCATCGTC 3113070

QY 343 CCCAATGTTGCCAAATATGCTCTCTGATACCATCTCTGTTATTGCCCACGAATCCTGTCTGAT 402

Db 3113069 GGCATGTTCATGGACAGCGGATTCGACGGCATCTTCTCTGTGGCGTCCAACCCAGTGGAT 3113010

QY 403 GTTTTGACCTATATTAGCTATAAGGCGCTCAGGGTTTCCACTAAGCAGAGTTATCGGGTCA 462

Db 3113009 ATCCTGACCTACGCAGTGTGGAATTTCTCCGGCTTGGATGGAACCGCGTATCGGCTCC 3112950

QY 463 GGTACAGTTCTGGATACCTGCTCGTTTAAATACATCTCCGAGAGCACTTCAAGATCTCA 522

Db 3112949 GGAATGTCTGGACTCCGCTCGATTCCGCTACATGCTGGCGAACTCTACGAAGTGCA 3112890

QY 523 TCGACAGCATCGATGCTCTGTAATTGGAGAACATGGTGATTCCGGTGTGCTCTG 582

Db 3112889 CCAAGCTCCGTCCACGCTTACATCATCGGGAACACCGGCGACACTGAACCTTCCAGTCTCTG 3112830

QY 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTCGGAAAAAGCCCAACCATATA 642

Db 3112829 TCCTCCGCGACCATCGCAGGCGTATCGCTTAGCCGAATGCTGACAAAGACCCAGAGCT- 3112771

QY 643 TTTGATCAGAAATCGTTCCATAGAATCTTTGAGCAACCGGAGACGCTGCTTACGATATC 702

Db 3112770 -----TGAGGGCCGCTTAGAGAAAATTTTCGAAGACACCCGCGACGCTGCCTATCACATT 3112716

QY 703 ATCAAGCGCAAAGGCTATATCTTCATATGGAATCGCAGCGGGATTACTTCGCATAGTAAAG 762

Db 3112715 ATCGACGCCAAGGGCTCCACTTCTACGGCATCGGCTGCTGCTCGCATCACCCGC 3112656

QY 763 GCGATTTTAGAGGATACAGGATCCACACTTACAGT---TTCAACCGTTGGTGATTATTTT 819

Db 3112655 GCAATCCTGCAGAACCAAGACGTTGCAGTCCCAGTCTCTGCACTGCTCCACGGTGAATAC 3112596

QY 820 GGGGTTGAACAAATTTGCTATFAAGCGTCCCTACCAAACTCAATAAAAGTGGGCTCATCAA 879

Db 3112595 GGTGAGGAAGACATCTACATCGGCACCCAGCTGTGGTGAACCGCGGAGGCATCCGCCGC 3112536

QY 880 GTGCTGAACTTTCACTCGATGAGAAGGAATAGAAATTGATGGAATAATCAGCTAGT 936

Db 3112535 GTTGTGAACTAGAAATCACCGACACGAGATGGAACGCTTCAAGCATTTCCGCAAT 3112479

RESULT 8

US-09-974-300-73

; Sequence 73, Application US/09974300

; Patent No. US20020146721A1

; GENERAL INFORMATION:

; APPLICANT: Berka, Randy M.

; APPLICANT: Clausen, Ib Groth

; TITLE OF INVENTION: Methods For Monitoring Multiple Gene

; TITLE OF INVENTION: Expression

; FILE REFERENCE: 10085.500-US

; CURRENT APPLICATION NUMBER: US/09/974,300

; CURRENT FILING DATE: 2001-10-05

; PRIOR APPLICATION NUMBER: 09/680,598

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/279,526

; PRIOR FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 8481

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 73

; LENGTH: 747

; TYPE: DNA

; ORGANISM: Bacillus licheniformis

US-09-974-300-73

Query Match 15.5%; Score 151; DB 11; Length 747;

Best Local Similarity 52.7%; Pred. No. 3.4e-35;

Matches 375; Conservative 0; Mismatches 330; Indels 6; Gaps 2;

QY 226 TATCCTGACTGCGCTGGCGGCCCATTTGTTATTGTGCATGTGGGATTAACCAAAAAAT 285

Db 2 TATGATGACTGCAAGAAGCTGATATTGTCGCTATGCGCGGAGCGAACCAAAAGCCG 61  
QY 286 GGACAAACAAGATGGATCTTGCTGCAAAAATGCCAACATTAATGCTGGAATCATCCCC 345  
Db 62 GGTGAAACCCGCTTGATCTAGTCGAAAAGAACTTGAAAATTTTAAAGGAATCATCGGA 121  
QY 346 AATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTGGCCAGCAATCCTGTCGATGT 405  
Db 122 GAAGTCATGGCCAGCGGTTTGGACGGCATCTTCTCGCTACAAAACCTGTCGACATC 181  
QY 406 TTGACCTATATTAGCTATAAGGCGTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCAGGT 465  
Db 182 TTGACTTATGCGACTTGGAATTCAGCGGCTTCCAAAAGAACGCGTCATCGGAAGCGGC 241  
QY 466 ACAGTTCTGGATACTGCTGCTGTTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCATCG 525  
Db 242 ACAACGCTTGACACTGCCCGTTTCCGCTACTTGCTCAGCGAGTATTTCGGGTGCTCGTGCC 301  
QY 526 GACAGCATCGATGCTGTGTAATTGGAGAACATGGTGATTCGGGTGCTGCTGCTGT 585  
Db 302 CACAATGCGCAGCGCTATATCATTTGGTGAACACGGAGATACGGAGCTGCTGCTGGAGC 361  
QY 586 CTTACCAACATCGACGGCATGAAGCTCCGGGATTACTCGGAAAAAGCAACACCATATTT 645  
Db 362 CATGCGAATATCGCGGAGTTCCTCGTCAGCG--ACTTGCTGAAAAGAAATGAGAAATAC 418  
QY 646 GATCAGAATGCGTTCATAGAACTCTTTGAGCAACGCGAGACGCTGCTTACGATATCATC 705  
Db 419 AAAGCGGAAGATCTAGATGAACCTGTTTGACAATGTAAAAACGCCGCTTACCACATTAT 478  
QY 706 AAGCGCAAGGCTATACCTCATATGGAATCGCAGCGGGATTACTTCGCATAGTAAAGCG 765  
Db 479 GAGAAAAAGCGCGACATACACTCGGAGTTGCCATGAGCCTCGCACGTATTACGAAAGCG 538  
QY 766 ATTTTAGAGATACAGGATCCACACTTACAGTTTCAAC---CGTTGGTGATTTTGGG 822  
Db 539 ATTTATCGAAATGAAGAAGCAATTTTGACCGTGAGCGCTCATCTTGATGGTGAATTGGA 598  
QY 823 GTTGAACAAATGCTATAAGCGTCCCTACCAAACTCAATAAAAGTGGGCTCATCAAGTG 882  
Db 599 GAAAATGATGTTTATATCGGAGTTCGGCGGTTGTCGGCCGATGCGGAGCGAGGAAATT 658  
QY 883 GCTGAACCTTCACTCGATGAGAAGGAATAGAAATTGATGGAAAAATCAGCT 933  
Db 659 GTTGAGCTCGATCTGAACGAAAAGGAAAAACAATTCAAGCATCGCGCAGGT 709

RESULT 9  
US-10-274-266-1  
; Sequence 1, Application US/10274266  
; Publication No. US20030059893A1  
; GENERAL INFORMATION:  
; APPLICANT: LADUNGA, Steven et al.  
; TITLE OF INVENTION: Isolated Human Secreted Proteins, Nucleic Acid Molecules Encoding  
; TITLE OF INVENTION: Secreted Proteins, And Uses Thereof  
; FILE REFERENCE: CL00839DIV  
; CURRENT APPLICATION NUMBER: US/10/274,266  
; CURRENT FILING DATE: 2002-10-21  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1146  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-274-266-1  
Query Match 14.0%; Score 136; DB 15; Length 1146;  
Best Local Similarity 49.7%; Pred. No. 1.5e-30;  
Matches 380; Conservative 0; Mismatches 375; Indels 9; Gaps 1;  
QY 52 AAAGTTGTGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC 111

Db 211 AAGTCTCCATCATAGGAACCTGGATCGGTGGGCATGGCCTGCGCTATCAGCATCTTATTA 270  
QY 112 AGCGCATCGTTTCCGAGATTGTCCTTATCGACGTGAACAAAGACAAAGCAGAGGGTGAA 171  
Db 271 AAAGGCTTGAGTGATGAACCTTGCCCTTGTTGGATCTTGATGAAGACAAACTGAAGGGTGAG 330  
QY 172 AGCATGGACTTAAACCAACGACGACCTTCAAATAACAAGGTC-----TCGAGCGGGT 222  
Db 331 ACGATGGATCTTCAACATGGCAGCCCTTTCACGAAATGCCAAATATGTTTGTAGCAAA 390  
QY 223 GATTATCCTGACTGCGTGGCGGCGCATTTGTTATTGTCACATGTCGCGATTAAACCAAAA 282  
Db 391 GATTACTTTGTCACAGCAAACTCCAACCTAGTGATTATCACAGCAGGTGCACGCCAAGAA 450  
QY 283 AATGGACAAAACAAGGATGGATCTTGTGCAAAAAAATGCCAACATTATGCTGGAAATCATC 342  
Db 451 AAGGAGAAACGCGCTTAATTTAGTCCAGCGAAATGTGGCCATCTTCAAGTTAATGATT 510  
QY 343 CCCAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTGCCACGAATCCTGTCGAT 402  
Db 511 TCCAGTATTGTCAGTACAGCCCCACTGCAAACTGATTATTGTTTCCAATCCAGTGGAT 570  
QY 403 GTTTTGACCTATATTAGCTATAAGCGGTCAAGGCTCAGGTTTCCACTAAGCAGAGTTATCGGCTCA 462  
Db 571 ATCTTAACTTATGTAGCTTGGAAGTTGAGTGCATTTCCCAAAAAACCGTATTATTGGAAGC 630  
QY 463 GGTACAGTTCTGGATACCTGCTGTTTTAAATACATCCTCGGAGACACTTCAAGATCTCA 522  
Db 631 GGCTGTAATCTGGATACCTGCTGTTTTCTGTTTCTTGATGGACAAAAGCTTGGTATCCAT 690  
QY 523 TCGGACAGCATCGATGCTGTGTAATTGGAGAACATGGTGATTTCGGGTGTCCTGCTGG 582  
Db 691 TCTGAAAGCTGCCATGGATGGATCCTCGGAGAGCATGGAGACTCAAGTGTTCTCTGTGG 750  
QY 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCAAACACATA 642  
Db 751 AGTGGAGTGAACATAGCTGGTGCTCCCTTTGAAGGATCTGAACTTGATATAGGAACATGAT 810  
QY 643 TTTGATCAGAAATCGTTTCCATAGAAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATATC 702  
Db 811 AAAGATCCTGAGCAATGGAAAAATGTCCACAAAGAAAGTGAAGTGAAGTGCCTATGAGATT 870  
QY 703 ATCAAGCGCAAAAGGCTATACCTTCAATGGAATCGCAGCGGGAATTAATTCGCATAGTAAAG 762  
Db 871 ATTAATAATGAAAGGTTATACCTTCTTGGGCCATTGGCCTATCTGTGGCCGATTTAACAGAA 930  
QY 763 GCGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAACCCGT 806  
Db 931 AGTATTTTGAAGAATCTTAGGAGAAATACATCCAGTTTCCACCAT 974

RESULT 10  
US-09-731-872-14  
; Sequence 14, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US3.REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 14  
; LENGTH: 1759  
; TYPE: DNA  
; ORGANISM: Homo sapiens

```

;
; NAME/KEY: CDS
; LOCATION: 101..1243
; NAME/KEY: sig_peptide
; LOCATION: 101..199
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 3.57142340200611
; OTHER INFORMATION: seq FLCGLGMCPRQA/TR
; US-09-731-872-14

```

Query Match		14.0%	Score 136;	DB 11;	Length 11759;
Best Local Similarity		49.7%	Pred. No. 1.9e-30;		
Matches	Conservative	0;	Mismatches 375;	Indels	Gaps 1;
QY	52	AAAGTTGTGGTAGTGGAGTGGGAAGTGTGGGTCTGCACAGCGTATACGTTGCTTCTC	111		
DB	311	AAGGTCTCCATCATAGGAACCTGGATCGGTGGGCATGCGCTCGCTATCAGCATCTTATTA	370		
QY	112	AGCGGCATCGTTTCCGAGATTGTCTTATCGACGTGAACAAAGACAAAGCAGAGSGTGAA	171		
DB	371	AAAGCTTGAGTGTGATGAACTTGCCCTTGTGGATCTTGATGAAGACAAACTGAAGGGTGAG	430		
QY	172	AGCATGGACTTAAACACGACGACACCTTCAAATACAAGTC-----TCGAGCGGGT	222		
DB	431	ACGATGGATCTTCAACATGGCAGCCCTTTACGAAATGCCAAATATTGTTTGTAGCAAA	490		
QY	223	GATTATCCTGACTGCGTGGCGCGGCCCATTTGTTATTGTACATGTGGATTAAACCAAAA	282		
DB	491	GATTACTTTGTACAGCAAACTTCAAACCTAGTGATTATCACAGCAGGTGCACGCCAAGAA	550		
QY	283	AATGGACAAACAAAGATGGATCTTGCTGCAAAAAATGCAAAACATTATGCTGGAATCATC	342		
DB	551	AAGGGAGAAACCGCCCTTAATTAGTCCAGCGAAATGTGCCCATCTTCAAGTTAATGATT	610		
QY	343	CCCAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTGCCACGAATCCTGTGCGAT	402		
DB	611	TCCAGTATTGTCCAGTACAGCCCCCACTGCAAACTGATTATTGTTTCCAATCCAGTGGAT	670		
QY	403	GTTTTGACCTATATTAGCTATAAGGCGTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCA	462		
DB	671	ATCTTAACTTATGTAGCTTGAAGTTGAGTGCACTTTCCAAAAACCGTATTATTGGAAGC	730		
QY	463	GGTACAGTTCTGGATACTGCTCGTTTAAATAACATCCTCGGAGAGCACTTCAAGATCTCA	522		
DB	731	GGCTGTAATCTGGATACTGCTCGTTTTCGTTTCTTGATGGACAAAAGCTTGGTATCCAT	790		
QY	523	TCGGACAGCATCGATGCTGTGTAATTGGAGAACATGGTGATTCGGGTGTCCCTGTCTGG	582		
DB	791	TCTGAAAGCTGCCATGGATGGATCCTCGGAGAGCATGGAGACTCAAGTGTTCTCTGTGTGG	850		
QY	583	TCTCTTACCACATCGACGSCATGAAGCTCCGGGATTACTCGGAAAAAGCCAACACATA	642		
DB	851	AGTGAGTGAACATAGTGTGTCCCTTTGAAGGATCTGAACTCTGTATATAGGAACCTGAT	910		
QY	643	TTTGATCAGAAATGCGTTCCATAGAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATATC	702		
DB	911	AAAGATCCTGAGCAATGGAATAATGTCCCAAAAGAGTGACTGCAACTGCCTATGAGATT	970		
QY	703	ATCAAGCGCAAGGCTTACTTTCATATGGAATCGCAGCGGGATTACTTCGCATAGTAAAG	762		
DB	971	ATTAAAAATGAAAGGTTTACTTCTTGGGCCATTGGCCCTATCTGTGGCCGATTTTAAACAGAA	1030		
QY	763	GCGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAACCGT	806		
DB	1031	AGTATTTTGAAGAATCTTTAGGAGAAATACATCCAGTTTCCACCAT	1074		

RESULT 11  
US-10-274-266-3

US 10 274 266 5  
; Sequence 3, Application US/10274266  
; Publication No. US2003005993A1  
; GENERAL INFORMATION:  
; APPLICANT: LADUNGA, Steven et al.

```

; TITLE OF INVENTION: Isolated Human Secreted Proteins, Nucleic Acid Molecules Encoding
; TITLE OF INVENTION: Secreted Proteins, And Uses Thereof
; FILE REFERENCE: CL000839DIV
; CURRENT APPLICATION NUMBER: US/10/274,266
; CURRENT FILING DATE: 2002-10-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3144
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-274-266-3

```

	Query Match	14.0%;	Score 136;	DB 15;	Length 3144;
	Best local Similarity	49.7%;	Pred. No. 2.6e-30;		
	Matches 380;	Conservative	0;	Mismatches 375;	Indels 9; Gaps 1;
QY	52	AAAGTTGTGTTAGTGGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC	111		
DB	1210	AAGGTCTCCATCATAGGAACCTGGATCGGTGGGCATGGCCTGCGCTATCAGCATCTTATTA	1269		
QY	112	AGCGGCATCGTTCCGAGATTGTCTTATCGACGTGAACAAAGACAAAGCAGAGGGTGAA	171		
DB	1270	AAAGGCTTGACTGATGAACCTTGCCCTTGTGGATCTTGTATGAAGACAAACTGAAGGGTGAG	1329		
QY	172	AGCATGGACTTAAACCACCGCAGCACCTTCAAATACAAGTC-----TCGAGCGGGT	222		
DB	1330	ACGATGGATCTTCAACATGGCAGCCCTTTCAGAAATGCCAAATATGTTTGTAGCAA	1389		
QY	223	GATTATCCTGACTGCGTGGCGGGGCCATTGTTATTGTACATGTGGGATTAACCAAAAA	282		
DB	1390	GATTACTTTGTACAGCAAACTCCAACTAGTGATTATCACAGCAGGTGCACGCCAAGAA	1449		
QY	283	AATGGACAAACAAGGATGGATCTTGCTGCAAAAAAATGCCAACATTATGCTGGAATCATC	342		
DB	1450	AAGGGAGAAACGCGCCTTAATTAGTCCAGCGAAATGTGGCCATCTTCAAGTTAATGATT	1509		
QY	343	CCCAATGTTGCCAAATATGCTCCTGATACCATCTGCTTATTGCCACGAATCCTGTGCGAT	402		
DB	1510	TCCAGTATTGTCCAGTACAGCCCCCACTGCAAACTGATTATTGTTTCCAATCCAGTGGAT	1569		
QY	403	GTTTTGACCTATATTAGCTATAAGGCGTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCA	462		
DB	1570	ATCTTAACTTATGTAGCTTGGAGTTGAGTGCAATTTCCCAAAACCGTATTATTGGAAAGC	1629		
QY	463	GGTACAGTCTGGATACCTGCTCGTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCA	522		
DB	1630	GGCTGTAATCTGGATACCTGCTCGTTTTCGTTTCTTGATTGGACAAAGCTTGGTATCCAT	1689		
QY	523	TCGGACAGCATCGATGCTGTGTAATTGGAGAACATGGTGATTCGGGTGTGCTGTCTGG	582		
DB	1690	TCTGAAAGCTGCCATGGATGGATCCTCGGAGAGCATGGAGACTCAAGTGTCTGTGTGG	1749		
QY	583	TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCAACACACATA	642		
DB	1750	AGTGGAGTGAACATAGCTGGTGTCCTTTGAAGGATCTGAACCTCTGATATAGGAACCTGAT	1809		
QY	643	TTTGTATCAGAAATGCGTTCCATAGAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATATC	702		
DB	1810	AAAGATCCTGAGCAATGGAAAAATGTCCCAAAAGAAGTGAAGTCACTGCTATGAGATT	1869		
QY	703	ATCAAGCGCAAGGCTTATCTTCATATGGAATCGCAGCGGGATTACTTCGCATAGTAAAG	762		
DB	1870	ATTAAATGAAGGTTATACCTTCTTGGGCCATTGGCCCTATCTGTGGCCGATTTAACAGAA	1929		
QY	763	GCGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAACCGT	806		
DB	1930	AGTATTTTGAAGAATCTTAGGAGAATACATCCAGTTTCCACCAT	1973		

RESULT 12  
US-09-938-842A-1938  
: Sequence 1938. Application US/09938842A



```
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1938
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1938
```

```
Query Match 12.7%; Score 123.6; DB 11; Length 1062;
Best Local Similarity 49.9%; Pred. No. 8.4e-27;
Matches 346; Conservative 0; Mismatches 339; Indels 9; Gaps 1;

QY 52 AAAGTTGTGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGGTATACGTTGCTTCTC 111
Db 124 AAAGTCTCCGTCTCGGAGTCGGAACGTCGGAATGGCCATAGCTCAAAACCATCCTCACT 183
QY 112 AGCGGCATCGTTTCCGAGATTGTCCTTATCGACGTGAACAAAGACAAAGCAGAGGGTGAA 171
Db 184 CAAGATCTCGCAGACGAGATCGCTCTCTGTTGACGCCAAACCCGATAAGCTTCGTGGCGAA 243
QY 172 AGCATGGACITAAACCAACGACGACACCTTC-----AAATACAGGCTCTCGAGCGGT 222
Db 244 ATGCTTGATCTCCAACACGCGCTGCTTCTCCTCCCTCGAATAGATCACCGCTTCTGTT 303
QY 223 GATTATCCTGACTGCGTGGCGGCGCAATTGTTATTGTCACATGTGGGATTAAACCAAAA 282
Db 304 GATTACGAAGTCACGCTGGATCTGATCTTGTATCGTCACTGCTGGTGTAGACAGAAT 363
QY 283 AATGGACAAACAAGGATGGATCTTGCTGCAAAAAAATGCCAACATATGCTGGAATATC 342
Db 364 CCAGGCGAGTCTAGACTCAATTTGCTTCAGAGGAACGTCGCTCTCTCCGTCATATCAIT 423
QY 343 CCCAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTGCCACGAATCCTGTGAT 402
Db 424 CCTCCACTCGTAAAGCTTCTCCTGATTTCTATATTGATCATTTGTTCTAATCCCGTTGAT 483
QY 403 GTTTTGACCTATATTAGCTATAAGGCGTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCA 462
Db 484 GTTTTGACTTACGTTGCTTGGAAACTCTCTGGCTTTCGGTGAATCGGCTTGGATCT 543
QY 463 GGTACAGTTCTGGATACTGCTCGTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCA 522
Db 544 GGTACTAATCTCGATTCTTCTCGGTTCCGATTCTTAATCGCAGATCATCTCGACGTTAAT 603
QY 523 TCGGACAGCATCGATGCCTGTGTAATTGGAGAACATCGTGAATTCGGGTGTGCTGTG 582
Db 604 GCTCAGGATGTACAGGCATTTATTGTGGGAGAGCATGGAGACAGCTCAGTGGCATTGTG 663
QY 583 TCTCTTACCAACATCGACGGCATGAAGTCCGGGATTACTCGGAAAAAGCCAAACACATA 642
Db 664 TCAAGCATTAGTGTGGAGGCATTCTCCTGCTTAAGCTTTTGGAGAGAAGAACAGATAGCT 723
QY 643 TTTGATCAGAAATCGTTCCATAGAACTTTTGAGCAACCGGAGAGCGCTGCTTACGATATC 702
Db 724 TACGAGAAACAAACTCTTGAGGATATCCACCAAGCTGTTGTTGGTAGCGCCTATGAAGTG 783
QY 703 ATCAAGCGCAAGGCTATACCTTCATATGGAATCG 736
```

```
Db 784 ATTGGACTCAAGGTTACACTTCTTGGGCCATTG 817

RESULT 13
US-09-971-361-1
; Sequence 1; Application US/09971361
; Patent No. US20020081677A1
; GENERAL INFORMATION:
; APPLICANT: Javed, Muhammad
; APPLICANT: Cusdin, Fiona
; APPLICANT: Milner, Paul
; APPLICANT: Green, Edward
; TITLE OF INVENTION: Ethanol Production
; FILE REFERENCE: 000487.00010
; CURRENT APPLICATION NUMBER: US/09/971,361
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/247,017
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: UK 0024554.8
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4913
; TYPE: DNA
; ORGANISM: Bacillus strain TN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4913)
; OTHER INFORMATION: n = A,T,C or G
US-09-971-361-1
```

```
Query Match 12.5%; Score 121.2; DB 10; Length 4913;
Best Local Similarity 54.1%; Pred. No. 1e-25;
Matches 293; Conservative 0; Mismatches 243; Indels 6; Gaps 2;

QY 377 TGCTTATTGCCACGAATCCTGTCGATGTTTGGACCTATATTAGCTATAAGCGTCAGGGT 436
Db 3758 TTCTTGTGGCAACGAACCCAGTGGATATTTTAACTGATGCTACTTGGAAATTTAGCGGGT 3817
QY 437 TTCCACTAAGCAGAGTTATCGGCTCAGGTACAGTTCTGGATACGTCGTTTTAAATACA 496
Db 3818 TACGAAAGAGCGGGTAATCGCTCAGGAACGATTCTTGATACAGCAAGATTCCGCTTCT 3877
QY 497 TCCTCGGAGAGCACTTCAAGATCTCATCGGACAGCATCGATGCTGTGTAATGGAGAAC 556
Db 3878 TGCTAAGTGAATATTTTCAAGTGGCTCCGACCAATGTACATGCTATATTATGGCGAGC 3937
QY 557 ATGCTGATTGGGTGTCCTGCTGTGCTCTCTTACCAACATCGACGGCATGAAGTCCGGG 616
Db 3938 ATGGGATACAGAGCTGCTGCTTGGAGCCATGCGGAATTTGGAAGCATTCAGTTGAGC 3997
QY 617 ATTACTCGGAAAAAGCCCAACACATATTGTATCAGAAATGCGTTCCATAGAAATCTTTGAGC 676
Db 3998 AAATATTGATGCAA---AACGATACTATAGAAAAGAGGATTAGACAATATCTTTGTTA 4054
QY 677 AAACGGGAGACGCTGTACGATATCATCAAGCGCAAGGCTATCTTCAATATGGAATCG 736
Db 4055 ATGTTCTGTGATGCGGCATATCAAAATCATTTAGAAAAGGGGCAACGTTATTACGGCATG 4114
QY 737 CAGCGGGATTACTTCGATAGTAAAGCGGATTTTAGAGGATACAGGATCCACACTTACAG 796
Db 4115 CAATGGGATTAGTCCGTATCACTCGTGTATTTTGCACAATGAAAATGCCATCTTAACCG 4174
QY 797 TTTCAACCGTTGGTGAT---ATTTGGGGTTGAACAAATTTGCTATAAGCGTCCCTACCA 853
Db 4175 TTTCTGCTCATTTGGACGGCCAATATGGCGAACGGAATGTTTATATTGGCGTGCCTGCCA 4234
QY 854 AACTCAATAAAAGTGGGCTCATCAAGTGGCTGAACCTTCACTCGATGAGAAGGAATAG 913
Db 4235 TTATCAACCGAAACGGTATTCTGTAAGTGATGGAATTGACGCTAAATGAAACAGAACAC 4294
```



QY 914 AA 915  
Db 4295 AA 4296

RESULT 14  
US-09-731-872-15  
; Sequence 15, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US3.REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 15  
; LENGTH: 1755  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 101..517  
; NAME/KEY: sig\_peptide  
; LOCATION: 101..199  
; OTHER INFORMATION: Von Heijne matrix  
; OTHER INFORMATION: score 3.57613483592743  
; OTHER INFORMATION: seq FLCIGMALCLRQA/TR  
US-09-731-872-15

Query Match 12.4%; Score 120.8; DB 11; Length 1755;  
Best Local Similarity 51.2%; Pred. No. 7.7e-26;  
Matches 392; Conservative 0; Mismatches 357; Indels 17; Gaps 4;

QY 52 AAAGTTGTGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC 111  
Db 311 AAGGTCTCCATCATAGGAACCTGGATCGGTGGGATGGCCTGCGCTATCAGCATCTTGTTA 370

QY 112 AGCGGCATCGTTCCGAGATTGCTCTTATCGACGTGAACAAAGACAAAGCAGAGGGTGAA 171  
Db 371 AAAGGCTTGAGTGATGAACCTGGCCCTTGCGATCTTGATGAAGACAAACTGAAGGGTGAG 430

QY 172 AGCATGGACTTAAACCCAGCAGCACCTTCAATPACAAGGTCTCGAGCGGT-----G 223  
Db 431 ACGATGGATCTTCAACATGGCAGCCCTTTCACGAAATGCCAATATTGTTGTAGCAAAG 490

QY 224 ATTATCCTGACTGCGTGGCGCGCCATTGTTATTGTACATGTGGGATTAAACCAAAAA 283  
Db 491 ATTACTTTGTACAGCAAACTCCAACTAGTATATACAGCAGGTGACGCCCAAGAA 550

QY 284 ATGGACAAACAAGGATGGATCTTGTGTCGCAAAAATGCCAACATTATGCTGGAATCATCC 343  
Db 551 AGGGAGAAACGGCCCTTAATTTAGTCCAGCGAAATGTGGCCATCTT-CAAGTAATGATTT 609

QY 344 CCAATGTTGCCAATATGCTCTGTATACCATCTGCTTATTGCCACGAATCCTGTCGATG 403  
Db 610 CCAGTATTGTCCAGTACAGCCCCCCTGCAAACTGATATTGTTTCCAACTCCAGTGGATA 669

QY 404 TTTTGACCTATATTAGCTATAAGGCGTCAAGGTTTCCACTAAGCAGAGTTATCGGCTCAG 463  
Db 670 TCTTAACCTTATGTAGCTTGAAGTTGAGTGCAATTTCCCAAAACCGTATTATTGGAAGCG 729

QY 464 GTACAGTTCTGGATACTGCTCGTTTTTAAATACATCCTCGGAGAGCACCTTCAAGATCTCAT 523  
Db 730 GCTGTAATCTGGATACTGCTCGTTTTTCGTTTCTTGATTTGGACAAAGCTTGGTATCCATT 789

QY 524 CGGACAGCATCGATCCCTGTGTAAATTGGAGAACATGGTGATTCGGGTGTGCTGTCTGGT 583  
Db 790 CTGAAAGCTGCCATGGATGGATCCTCGGAGAGCATGGAGACTCAAGTGTTCCTGTGTGGA 849

QY 584 CTCTTACCAACATCGACGGCATGAAGCTCCGGAT---TACTCGAAAAAGCCCAACCACA 640  
Db 850 GTGGAGTGAAACATAGCTGGTGTCCCTTTGAAGGATCTGAAGTCTGATATAGGAACCTGATA 909

QY 641 TATTTGATCAGAAATCGTTTCCATAGAATCTTTGAGCAACCGGAGACGCTGCTTACGATA 700  
Db 910 AAGATCCTGAGCAGG-----AAAAATGTCCACAAAGAGTGAAGTCAACTGCCTATGAGA 964

QY 701 TCATCAAGCGCAAGGCTATATCTTATATGGAATCGCAGCGGGAATTACTTCCATAGTAA 760  
Db 965 TTATTAAAAATGAAAGGTTATATCTTCTTGGGCCATTGGCCTATCTGTGGCCGATTTAACAG 1024

QY 761 AGGCGATTTTAGAGGATACAGGATCCACACTTACAGTTTCAACCGT 806  
Db 1025 AAAGTATTTTGAAGAATCTTAGGAGAATACATCCAGTTTCCACCAT 1070

RESULT 15  
US-09-974-298-74  
; Sequence 74, Application US/09974298  
; Patent No. US20020156263A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Huei-Mei  
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER  
; FILE REFERENCE: PA-0037 P  
; CURRENT APPLICATION NUMBER: US/09/974,298  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: 60/238,331  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 194  
; SOFTWARE: PERL Program  
; SEQ ID NO 74  
; LENGTH: 1172  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020156263A1 3365683CB1  
US-09-974-298-74

Query Match 11.3%; Score 109.8; DB 11; Length 1172;  
Best Local Similarity 47.7%; Pred. No. 1.3e-22;  
Matches 361; Conservative 0; Mismatches 387; Indels 9; Gaps 1;

QY 61 GTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTCAGCGGCATC 120  
Db 191 GTAGTGGGTGTGGACAAGTTGGTATGGCGGTGTGTATCAGCATCTGGGAAAGTCTCTG 250

QY 121 GTTTCGAGATTGTCTTATCGAGCTGAACAAAGACAAAGCAGAGGGTGAAAGCATGGAC 180  
Db 251 GCTGATGAACCTGCTCTTGTGGATGTTTGGAAAGATAAGCTTAAAGGAGAAATGATGGAT 310

QY 181 TTAAACCAAG-----CAGCACCTTCAAATACAAGGTCTCGAGCGGGTGATTATCCT 231  
Db 311 CTGCAGCATGGAGCTTATTTCTTCAGACACCTAAATTTGTGCAGATAAAGATTATCT 370

QY 232 GACTGCGTGGCGCGGCCATTGTTATTGTACATGTGGGATTAAACCAAAAATGGACAA 291  
Db 371 GTGACCGCCAATTCTAAGATTGTAGTGGTAACTGCAGGAGTCCGTCAGCAAGAAGGGGAG 430

QY 292 ACAAGGATGGATCTTGTCTGCAAAAAATGCCAACATTATGCTGGAATCATCCCCAATGTT 351  
Db 431 AGTCGGCTCAATCTGGTGCAGAGAAATGTTAATGTCTTCAAAATTCATTATTCCTCAGATC 490

QY 352 GCCAATATGCTCCTGATACCATCTGCTTATTGCCACGAATCCTGTGATGTTTTGACC 411  
Db 491 GTCAAGTACAGTCTGATTTGTCATCAATTTGTGTTTCCAAACCCAGTGGACATTTCTTACG 550

QY 412 TATATTAGCTATAAGGCGTCAAGGTTTCCACTAAGCAGAGTTATCGGCTCAGGTACAGTT 471





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: July 31, 2003, 08:17:36 ; Search time 1718 Seconds  
(without alignments)  
9162.988 Million cell updates/sec

Title: US-09-992-430B-21  
Perfect score: 972  
Sequence: 1 atgttccaagatacaaaagtc.....ttgagcatctggagatcaat 972

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	152	15.6	968	17 CNS06Y1R	AL420469 T3 end of
2	135	13.9	936	14 BQ963532	BQ963532 AGENCOURT
3	134.8	13.9	1058	13 BI410782	BI410782 602963554
C 4	133.4	13.7	2603	17 BH770979	BH770979 LLMGtag70
5	131.6	13.5	913	14 BQ895234	BQ895234 AGENCOURT
6	131.6	13.5	918	14 BQ955812	BQ955812 AGENCOURT

7	131.6	13.5	1032	13 BI410482	BI410482 602963152
8	131.4	13.5	891	13 BI905403	BI905403 603167455
9	131.4	13.5	1097	9 AU090981	AU090981 AU090981
10	130.8	13.5	951	14 BQ895574	BQ895574 AGENCOURT
11	130.6	13.4	903	14 BQ884746	BQ884746 AGENCOURT
12	130	13.4	959	14 BQ886472	BQ886472 AGENCOURT
C 13	129.8	13.4	645	17 CNS06YKH	AL421143 T3 end of
14	129.8	13.4	926	14 BQ899735	BQ899735 AGENCOURT
15	129.6	13.3	924	13 BG916902	BG916902 602816178
16	128.8	13.3	894	13 BI904831	BI904831 603168808
17	128.8	13.3	1076	14 BQ918069	BQ918069 AGENCOURT
18	127.4	13.1	902	14 BQ886973	BQ886973 AGENCOURT
19	126.6	13.0	1129	13 BM464639	BM464639 AGENCOURT
20	126	13.0	824	13 BI905607	BI905607 603167795
21	126	13.0	1082	13 BM476010	BM476010 AGENCOURT
22	125.8	12.9	1108	13 BM462026	BM462026 AGENCOURT
23	125.6	12.9	710	14 BQ571108	BQ571108 UI-M-PB0-
24	125.6	12.9	818	13 BG915740	BG915740 602814260
25	125.4	12.9	854	13 BI410650	BI410650 602963448
26	125.2	12.9	944	9 AL559924	AL559924 AL559924
27	125	12.9	920	9 AL532872	AL532872 AL532872
28	125	12.9	947	9 AL543215	AL543215 AL543215
29	125	12.9	964	9 AL545000	AL545000 AL545000
30	125	12.9	1020	9 AL554514	AL554514 AL554514
31	125	12.9	1021	9 AL553512	AL553512 AL553512
32	125	12.9	1023	9 AL514320	AL514320 AL514320
33	125	12.9	1023	9 AL550051	AL550051 AL550051
34	124.6	12.8	949	14 BQ650107	BQ650107 AGENCOURT
35	124.6	12.8	965	9 AL552670	AL552670 AL552670
36	124.6	12.8	1017	9 AL559984	AL559984 AL559984
37	124.2	12.8	963	9 AL559795	AL559795 AL559795
38	124	12.8	999	9 AL541618	AL541618 AL541618
39	124	12.8	1059	13 BM450615	BM450615 AGENCOURT
40	123.8	12.7	958	9 AL513680	AL513680 AL513680
41	123.6	12.7	741	14 BQ870980	BQ870980 QGI10J04.
42	123.6	12.7	842	9 AL541794	AL541794 AL541794
43	123.4	12.7	967	9 AL549487	AL549487 AL549487
44	123.4	12.7	967	14 BQ956047	BQ956047 AGENCOURT
45	123.2	12.7	872	14 BQ212714	BQ212714 AGENCOURT

ALIGNMENTS

RESULT 1	CNS06Y1R/c	CNS06Y1R	968 bp	DNA	linear	GSS 06-JUL-2001
LOCUS	T3 end of clone AY0AA005E09 of library AY0AA from strain CBS 6340					
DEFINITION	of Kluyveromyces thermotolerans, genomic survey sequence.					
ACCESSION	AL420469					
VERSION	AL420469.1	GI:12203655				
KEYWORDS	GSS.					
SOURCE	Kluyveromyces thermotolerans.					
ORGANISM	Kluyveromyces thermotolerans					
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetales; Kluyveromyces					
AUTHORS	1 (bases 1 to 968) Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,W., Tekaiia,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissbach,J.					
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies					
JOURNAL	FEBS Lett. 487 (1), 3-12 (2000)					
MEDLINE	20584711					
PUBMED	11152876					
REFERENCE	2 (bases 1 to 968)					
AUTHORS	Malpertuy,A., Llorente,B., Blandin,G., Artiguenave,F., Wincker,P. and Dujon,B.					
TITLE	Genomic exploration of the hemiascomycetous yeasts: 10. Kluyveromyces thermotolerans					



JOURNAL FEBS Lett. 487 (1), 61-65 (2000)  
MEDLINE 20584720  
PUBMED 11152885  
REFERENCE 3 (bases 1 to 968)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,  
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT This GSS is part of a random genomic sequencing program of thirteen  
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces  
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,  
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces  
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia  
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,  
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to  
5 kb were prepared and both extremities were sequenced. See  
keywords for description of this sequence and for the sequence of  
the other extremity of this insert.

FEATURES Location/Qualifiers  
source  
1. .968  
/organism="Kluyveromyces thermotolerans"  
/strain="CBS 6340"  
/db\_xref="taxon:4916"  
/clone="AY0AA005E09"  
/clone\_lib="AY0AA"  
/note="end : T3"  
misc\_feature complement(<4. .>357)  
/note="similar to Saccharomyces cerevisiae ORF YDL078c [MDH3 ; malate dehydrogenase, peroxisomal]"  
evidence=not\_experimental

BASE COUNT 279 a 199 c 218 g 267 t 5 others  
ORIGIN  
Query Match 15.6%; Score 152; DB 17; Length 968;  
Best Local Similarity 66.2%; Pred. No. 1.5e-33;  
Matches 237; Conservative 0; Mismatches 115; Indels 6; Gaps 1;

Qy 38 ATGCCAAAACAGTAAAGTTGTGGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGT 97  
Db 368 AAGTGAACCCCTGTCAAGGTGTCTGTTATTGGTGTGGGAGTGTGGGCTCTACTACGGCCT 309  
Qy 98 ATACGTTGCTTCTCAGCGGCATCGTTTCCGAGATTGCTCTTATCGACGTGAACAAAGACA 157  
Db 308 ACACCTGCTATTAGCGAGATGATCTCGGAAGTGTGTTGATTGATATCAACACACATA 249  
Qy 158 AAGCAGAGGCTGAAGCATGGACTTAAACCCAGCAGCACCTTCAAAATACAAGGTCTCGAG 217  
Db 248 AAGCAGAGGGCGAAGCATGGATTAAACCATGCTGCACCTTCCACACAGGTTCTGTTG 189  
Qy 218 CG-----GGTGATTATCCTGACTGCGCTGGCGGCCCATTTGTTATTGTACATGTGGGA 271  
Db 188 TGTATGTTGGAGATTATAGCGACTGCGCCGACGCTGCGATTGTGATCATACAGGTGGCG 129  
Qy 272 TTAACCAAAAAAATGACAAACAAGGATGGATCTTGTGCAAAAAAATGCCAACATTATGC 331  
Db 128 CCAATCAAAAAACCTGTCAAAACCTAGATGGATTAGCAGCAACAAACGCAAGAATCTTAC 69  
Qy 332 TGGAAATCATCCCAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTGCCAC 389  
Db 68 AAGGTATTATTCCCAAGATAGTTGAGTACGCCCAAAAAACTATTCTTCTCATTGCCAC 11

RESULT 2  
BQ963532 936 bp mRNA linear EST 21-AUG-2002  
LOCUS  
DEFINITION AGENCOURT\_10050420 NIH\_MGC\_134 Mus musculus cDNA clone  
IMAGE:6514321 5', mRNA sequence.  
ACCESSION BQ963532  
VERSION BQ963532.1 GI:22379010  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 936)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Dr. David Rowe  
cDNA Library Preparation: Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM14089 row: h column: 02  
High quality sequence stop: 683.  
FEATURES Location/Qualifiers  
source  
1. .936  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6514321"  
/clone\_lib="NIH MGC 134"  
/tissue\_type="undifferentiated limb"  
/lab\_host="DH10B (phage-resistant)"  
/note="Vector: pCMV-SPORT6.1.cdb; Site 1: EcoRV; Site 2:  
NotI; Cloned unidirectionally. Primer: Oligo dt. Average  
insert size 1.7 kb. Constructed by ResGen, Invitrogen  
Corp. Note: this is a NIH\_MGC Library."

BASE COUNT 223 a 237 c 264 g 212 t  
ORIGIN  
Query Match 13.9%; Score 135; DB 14; Length 936;  
Best Local Similarity 51.1%; Pred. No. 1.5e-28;  
Matches 349; Conservative 0; Mismatches 325; Indels 9; Gaps 1;  
Qy 52 AAAGTTGTGGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC 111  
Db 173 AAGATTACAGTTGTTGGGGTGTGGTCTGTTGGCATGGCTTGTCCATCAGTATCTTAATG 232  
Qy 112 AGCGGCATCGTTTCCGAGATTGTCTTATCGACGTGAACAAAGACAAAGCAGAGGTGAA 171  
Db 233 AAGCACTTGGCGGATGAGCTTGCCCTTGTGACGTATGGAAGACAAACTCAAGGGCGAG 292  
Qy 172 AGCATGGACTTAAACACCGCAGCACCTTCAAAATACAAGGTCTCGAGCGGTGATTATCCT 231  
Db 293 ATGATGGATCTCCAGCATGGCAGCCTCTTCTCTTAAACACCAAAAATGTCTCCAGCAA 352  
Qy 232 GACTGCGCTGGCGGCCA-----TTGTTATTGTACATGTGGGATTAAACCAAAA 282  
Db 353 GACTACTGTGTAAGTGGCAACTCCAAAGCTGGTCAATTATCACCGGGGGCCCGTCAGCAA 412  
Qy 283 AATGCACAAACAAGGATGGATCTTGTGCAAAAAATGCCAACATTATGCTGGAATCATC 342  
Db 413 GAGGGGAGAGCCGGCTCAACCTGGTCCAGCGAAACGTGAACATCTTCAAGTTCATCAT 472  
Qy 343 CCCAATGTTGCCAAATATGCTCTGTATACCATCCTGCTTATTTCCACGAATCCTGTCGAT 402  
Db 473 CCCAACATTGTCAAGTACAGTCCACACTGCAAGCTGCTGATGCTCTCCAATCCAGTGGAT 532  
Qy 403 GTTTTGACCTATATTAGCTATAAGGCGTCAAGGGTTTCCACTAAGCAGAGTTATCGGCTCA 462  
Db 533 ATCTTGACCTACGTGGCTTGGAAAATCAGTGGCTTTCCCAAAAAACCGAGTAATTGGAAGT 592  
Qy 463 GGTACAGTTCTGGATACCTGCTCTTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCA 522  
Db 593 GGTGCAATCTGGATTTCAGCGCGGTTCGGTTACCTGTATGGGAGAGAGGCTGGGGTTTCA 652  
Qy 523 TCGGACAGCATCGATGCTCTGTAATTGGAGAACATGTTGGGTGTGCGCTGTCTGG 582  
Db 653 GCGCTGAGCTGTACCGGCTGGTCTCTGGGAGAACATGGCGACTCCAGTGTGCTGTGG 712  
Qy 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCACACATA 642

Db 713 AGTGGTGAATGTTGCCGGCTCTCCCTGAAGTCTCTTAACCCAGAACTGGGCACTGAC 772  
QY 643 TTTGATCAGATCGGTTCCATAGAAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATATC 702  
Db 773 GCAGACAAGGAGCAGTGGGAAGGAGTTTCAAGAAGAGTGGTGACAGTGCCTACGAGGTG 832  
QY 703 ATCAAGCGCAAGGCTATCTTC 725  
Db 833 ATCAAGCTGAAAGGTTACACATC 855

RESULT 3  
BI410782  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BI410782 1058 bp mRNA linear EST 14-AUG-2001  
602963554F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5119236 5',  
mRNA sequence.  
BI410782  
BI410782.1 GI:15171705  
EST.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1058)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM1290 row: k column: 13  
High quality sequence start: 30  
High quality sequence stop: 917.  
Location/Qualifiers  
1. .1058  
/organism="Mus musculus"  
/strain="CZECH II"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5119236"  
/clone\_lib="NCI\_CGAP\_Lu33"  
/tissue\_type="pooled lung tumors"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a  
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st  
strand cDNA was prepared from mRNA obtained from pooled  
lung tumors with a Not I - oligo(dt) primer [5',  
TGTTACCAATCTGAAGTGGAGCGCGCCTCTGTTTTTTTTTTTTTTT 3'].  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not  
I and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 265 a 258 c 293 g 242 t  
ORIGIN  
Query Match 13.9%; Score 134.8; DB 13; Length 1058;  
Best Local Similarity 50.1%; Pred. No. 1.9e-28;  
Matches 368; Conservative 0; Mismatches 357; Indels 9; Gaps 1;  
QY 52 AAAGTTGTGTAGTGGAGTGGGAAGTGTGGTCTGCCACAGCGTATACGTTGCTTCTC 111  
Db 152 AAGATTACAGTTGTGGGTTGGTGTCTGTGGCATGGCTTGTGCCATCAGTATCTTAATG 211  
QY 112 AGCGGCATCGTTTCCGAGATTGTCCTATCGACGTTGAACAAAGACAAAGCAGAGGTTGAA 171

Db 212 AAGGACTTGGCGGATGAGCTTGCCCTTGTGTGACGTATGGAAGACAAACTCAAGGGCGAG 271  
QY 172 AGCATGGACTTAAACCCAGCAGCACCTTCAAATACAAGGTCTCGAGCGGTGATTATCCT 231  
Db 272 ATGATGGATCTCCAGCATGGCAGCCTCTTCTTAAACACCAAAAATTTGTCTCCAGCAA 331  
QY 232 GACTGCGCTGGCGGCCA-----TTGTTATTGTCAATGTGGGATTAAACCAAAA 282  
Db 332 GACTACTGTGTAAGTGCAGAACTCCAAGCTGGTCAATTATCACCGGGGGCCCGTCAGCAA 391  
QY 283 AATGACAAACAAGGATGGATCTTGTCTGCAAAAAATGCCAACATTATGCTGGAATCATC 342  
Db 392 GAGGGGAGAGCCGGCTCAACCTGGTCCAGCGAAACGTGAACATCTTCAAGTTCATCA 451  
QY 343 CCCAATGTTGCCAAATATGCTCTCTGATACCATCTGCTTATTGCCACGAATCCTGTCGAT 402  
Db 452 CCCAACATTTGTCAGTACAGTCCACACTGCAAGCTGCTGATGCTCTCCAATCCAGTGGAT 511  
QY 403 GTTTTGACCTATATTAGCTATAAGGCGTCAGGGTTTCCACTAAGCAGAGATTATCGGCTCA 462  
Db 512 ATCTTGACCTACGTGGCTTGAAAAATCAGTGGCTTTCCCAAAAAACCGAGTAATTGGAAGT 571  
QY 463 GGTACAGTTCTGGATACGTCTCGTTTTTAAATACATCTTCGAGAGACACTTCAAGATCTCA 522  
Db 572 GGTGCAATCTGGATTACGCGCGGTTCCGTTACCTGATGGGAGAGAGGCTGGGGTTTCA 631  
QY 523 TCGGACAGCATCGATGCCCTGTGTAATTGGAGAACATGTTGATTCGGGTGTGCCTGTCTGG 582  
Db 632 GCGCTGAGCTGTACGCGCTGGTCTCTGGGAGAACATGGCGACTCCAGTGTGCCTGTGTGG 691  
QY 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCCAACACATA 642  
Db 692 AGTGGTGAATGTTGCCGGCGTCTCCCTGAAGTCTCTTAACCCAGAACTGGGCACTGAC 751  
QY 643 TTTGATCAGAATCGGTTCCATAGAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATATC 702  
Db 752 GCAGACAAGGAGCAGTGGAGAGGTTTCAAAAGCAGGTGTGTGACAGTGCCTACGAGGTG 811  
QY 703 ATCAAGCGCAAGGCTATATTCTATATGGAATCGCAGCGGATTACTTCGCATAGTAAAG 762  
Db 812 ATCAAGCTGAAAGGTTACACATCTCTGGGCCATGGCCTCTCTGTGGCAAGACTTGGTGAAA 871  
QY 763 GCGATTTTAGAGGA 776  
Db 872 GCATAATGAAAGAA 885

RESULT 4  
BH770979/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BH770979 2603 bp DNA linear GSS 01-MAY-2002  
LLMGtag704 MG1363 Random Sequence Tag Library Lactococcus lactis  
subsp. cremoris genomic, DNA sequence.  
BH770979  
BH770979.1 GI:20373936  
GSS.  
Lactococcus lactis subsp. cremoris.  
Lactococcus lactis subsp. cremoris  
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
Lactococcus.  
1 (bases 1 to 2603)  
Bolotin,A., Ehrlich,S.D. and Sorokin,A.  
Studies of genomes of dairy bacteria Lactococcus lactis  
Sci. Aliments, (2002) In press  
Contact: Sorokin A  
Genetique Microbienne  
INRA  
CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France  
Tel: 33 1 34 65 25 16  
Fax: 33 1 34 65 25 21  
Email: [sorokine@jouy.inra.fr](mailto:sorokine@jouy.inra.fr)  
best homologue in strain IL1403 is ldh (100%)  
Class: shotgun  
High quality sequence start: 30

High quality sequence stop: 2575.  
Location/Qualifiers  
1. .2603  
/organism="Lactococcus lactis subsp. cremoris"  
/strain="MGL1363"  
/db\_xref="taxon:1359"  
/clone\_lib="MGL1363 Random Sequence Tag Library"  
/note="Vector: pSGMU2; Site 1: SmaI; Library of  
chromosomal fragments of L.lactis strain MGL1363 was  
prepared by partial AluI digestion or by sonication."  
BASE COUNT 743 a 550 c 496 g 814 t  
ORIGIN

Query Match 13.7%; Score 133.4; DB 17; Length 2603;  
Best Local Similarity 50.2%; Pred. No. 7.4e-28;  
Matches 388; Conservative 0; Mismatches 376; Indels 9; Gaps 2;

QY 35 CTGATGCCAAACACGTAAGTTGTGGTAGTGGGAGTGGGAAGTGTGGGTCTGCCACAG 94  
|||||  
Db 1895 CTGATAAACACGTAAGAAAGTTATCCCTGTGGTGACGGTGTGAGGTTTCATCATACG 1836

QY 95 CGTATACGTTGCTCTCAGCGGCATCGTTTCGAGATTGTCCTTATCGACGTGAACAAAG 154  
|||||  
Db 1835 CTTTGGCCCTGTAAACCAAGGAATTGCACAGAATTAGGTATTGTGACCTTTTAAAG 1776

QY 155 ACAAAGCAGAGGGTGAAGCATGGACTTAAACCA-----CGCAGCACCTTCAAATACAA 208  
|||||  
Db 1775 AAAAACTCAAGGGGATGCAGAAGACCTTTCTCATGCCTTGGCATTATACATCACCTAAA 1716

QY 209 GGTCTCGAGCGGTGATTATCTGACTGCGCTGGCGGGCCATTGTTATTGTACATGIG 268  
|||||  
Db 1715 AGATTACTCTGCAGACTACTCTGATGCAAGCAGCGTGACCTCGTTGTTGACTTCTG 1656

QY 269 GGATTAACCAAAAAATGGACAAACAGGATGGATCTTGTGTCACAAAAATGCCAACATTA 328  
|||||  
Db 1655 GTGCTCCACAAAAACAGGTGAACACTCGTCTTGACCTTGTGAAAAAAATCTTCGTATTA 1596

QY 329 TGCTGGAATCATCCCCAATGTTGCCAAATATGCTCCTGATACCATCTCGTTATTGCCA 388  
|||||  
Db 1595 CTAAAGATGTTGTAACATAAATGTTGCTTCAGGATCAAAGGAATCTCCTCGTTGCTG 1536

QY 389 CGAATCTGTGATGTTTGTGACCTATATTAGTATAAAGCGTCAAGGTTTCCACTAAGCA 448  
|||||  
Db 1535 CTAACCCAGTTGACATCTTGACATACGCAACTTGGAAATTTCTGTTTCCCTAAAAACC 1476

QY 449 GAGTTATCGGCTCAGTACAGTTCTGGATACGTCTGTTTAAATACATCCTCGGAGAGC 508  
|||||  
Db 1475 GTGTTGATGTTGAGTACTTCACTTGATACGACGTTCCGTCAGCATTTGGCTGAAA 1416

QY 509 ACTTCAAGATCTCATCGGACAGCATCGATGCTGTGTAATTGGAGAACATGTTGATTCG 568  
|||||  
Db 1415 AAGTTGACGTTGATGCTCGTTCAATCCACGCATACATCATGGTGAACACGGTGACTCAG 1356

QY 569 GTGTGCCTGTCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAA 628  
|||||  
Db 1355 AATTGCTGTTGGTACACGCTAACGTTGCTGGTTAAATTT---GGAACAATGGTTCC 1299

QY 629 AAGCCAAACCATATTGATCAGAATGCGTTCCATAGAAATCTTTGAGCAACGCGAGACG 688  
|||||  
Db 1298 AAGAAAATGACTACTTAACGAAGCAGAAATCGTTGAATGTTTGAATGTTGATGATG 1239

QY 689 CTGCTTACGATATCATCAAGCGCAAGGCTATATCTCATATGGAATCGCAGCGGATTAC 748  
|||||  
Db 1238 CAGCTTACTCAATCATCGCTAAAAAAGGTGCAACATTTACGGTGTGGCTGTAGCCCTTG 1179

QY 749 TTCGCATAGTAAAGCGGATTTTAGAGGATACAGGATCCACACTTACAGTTTCA 801  
|||||  
Db 1178 CTCGTATTACTAAAGCAATCTTGTGATGATGAACATGCAGTACTTCTCTGTATCA 1126

RESULT 5  
BQ895234 913 bp mRNA linear EST 16-AUG-2002  
LOCUS

DEFINITION AGENCOURT\_8752520 NIH\_MGC\_130 Mus musculus cDNA clone IMAGE:6333286  
5', mRNA sequence.  
ACCESSION BQ895234  
VERSION BQ895234.1 GI:22287248  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 913)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,  
Ph.D.  
cDNA Library Preparation: ResGen, Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM13792 row: d column: 23  
High quality sequence stop: 740.  
FEATURES  
Location/Qualifiers  
1. .913  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6333286"  
/clone\_lib="NIH\_MGC\_130"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: otcysts; Vector: pCMV-SPORT6.1.cdb;  
Site\_1: EcoRV; Site\_2: NotI; Cloned unidirectionally.  
Primer: Oligo dt. Average insert size 1.95 kb.  
Constructed by ResGen, Invitrogen Corp. Note: this is a  
NIH MGC Library."  
BASE COUNT 219 a 232 c 255 g 206 t 1 others  
ORIGIN

Query Match 13.5%; Score 131.6; DB 14; Length 913;  
Best Local Similarity 50.6%; Pred. No. 1.5e-27;  
Matches 351; Conservative 0; Mismatches 334; Indels 9; Gaps 1;

QY 52 AAAGTTGTGTAGTGGGAGTGGGAAGTGTGGTCTGCCACAGGTATACGTTGCTTC 111  
|||||  
Db 153 AAGATTACAGTTGTTGGGTTGGTGTGGTGTGGCATGGCTTGTCCATCAGTATCTTAATG 212

QY 112 AGCGGCATCGTTTCCGAGATTGCTCTTATCGACGTGAACAAAGCAGAGGGTGAA 171  
|||||  
Db 213 AAGGACTTGGCGGATGAGCTTGCCCTTGTGACGTATGGAAGACAAACTCAAGGGCGAG 272

QY 172 AGCATGGACTTAAACCACGACGACCTTCAAATACAAGTCTCGAGCGGTGATTATCCT 231  
|||||  
Db 273 ATGATGGATCTCCAGCATGGCAGCTCTTCTTAAACACCAAAAAATGTCTCCAGCAA 332

QY 232 GACTGCGCTGGCGGCCA-----TTGTTATTGTACATGTGGGATTAAACCAAAA 282  
|||||  
Db 333 GACTACTGTAACTGCGAACTCCAAGCTGGTCAATTATCACCGGGGGCCCGTCAGCAA 392

QY 283 AATGCACAAACAGGATGATCTTGTCTGCAAAAAATGCCAACATTATGCTGGAATCATC 342  
|||||  
Db 393 GAGGGGAGAGCGGGCTCAACCTGTCAGCGGAAACGCTGAACATCTTCAAGTTTCATT 452

QY 343 CCCAATGTTCCAAATATGCTCTGTATACCATCTGCTTTATGCCACGAATCCTGTCAT 402  
|||||  
Db 453 CCCAATGTTCAAGTACATCCACACTGCAAGCTGCTGATGCTCTCCAATCCAGTGGAT 512

QY 403 GTTTTGACCTATATTAGCTATAGGCGTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCA 462  
|||||  
Db 513 ATCTTGACCTACGTGGCTTGGAAATCAGTGGCTTTCCCAAAACCGAGTAATTGGAAGT 572

QY 463 GGTACAGTTCTGGATACGTGCTCGTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCA 522



Db 573 GGTGCAATCTGGATTGAGCGCGGTTCCGTTACCTGATGGGAGAGAGGCTGGGGTTTCAC 632

Qy 523 TCGGACAGCATCGATCGCTGTGTAATTGGAGAAACATGGTGATTCGGGTGCTGTCTGG 582

Db 633 GCGCTGAGCTGTACGCGCTGGTCTCTGGGAGAACATGGCGACTCCAGTGTGCTGTGTGG 692

Qy 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGGAAAGCCAAACACATA 642

Db 693 AGTGGTGTGAATGTTCCCGCGTCTCCCTGAAGTCTCTTAACCCAGAACTGGGCACTGAC 752

Qy 643 TTTGATCAGAATGCGTTCCATAGAATCTTTGAGCAACCGGAGACGCTGCTTACGATATC 702

Db 753 GCAGACAAGGAGCAGTGAAGAGGTTTCAAGCAGGTGGTGACAGTGCCTACGAGGTG 812

Qy 703 ATCAAGCGGAAAGGCTATATCTTCATATGGAATCG 736

Db 813 ATCAAGCTGAAAGGTTACACATCTCTGGGCCATTG 846

RESULT 6

BQ955812

LOCUS

DEFINITION

AGENCOURT\_8752602 NIH\_MGC\_130 Mus musculus cDNA clone IMAGE:6335643

5', mRNA sequence.

BQ955812

VERSION

BQ955812.1 GI:22371290

KEYWORDS

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 918)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.

cDNA Library Preparation: ResGen, Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM13798 row: g column: 04

High quality sequence stop: 716.

Location/Qualifiers

1..918

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:6335643"

/clone\_lib="NIH\_MGC\_130"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: otocysts; Vector: pCMV-SPORT6.1.ccdB; Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH\_MGC Library."

BASE COUNT 222 a 223 c 261 g 212 t

ORIGIN

Query Match 13.5%; Score 131.6; DB 14; Length 918;

Best Local Similarity 50.6%; Pred. No. 1.5e-27;

Matches 351; Conservative 0; Mismatches 334; Indels 9; Gaps 1;

Qy 52 AAAGTTGTGTAGTGGGAGTGGGAAGTGTGGTCTGCCACAGCGTATACGTTGCTTCTC 111

Db 101 AAGATTACAGTTGTTGGGTTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 160

Qy 112 AGCGGCATCGTTCCGAGATTGTCCTTATCGACGTTGAACAAAGACAAAGCAGAGGTTAA 171

Db 161 AAGGACTTGGCGGATGAGCTTGCCCTTGTGACGTATGGAAGACAAACTCAAGGGCGAG 220

Qy 172 AGCATGGACTTAAACCACGAGGACCTTTCAAATACAAGGTCTCGAGCGGTGATTATCCT 231

Db 221 ATGATGGATCTCCAGCATGGCAGCCTCTTCTTAAACACCAAAAATTTGTCTCCAGCAA 280

Qy 232 GACTGCGCTGGCGGCCA-----TTGTTATTGTCACTGTGGGATTAAACCAAAA 282

Db 281 GACTACTGTGTAAGTGCAGAACTCCAAGCTGGTCAATTATCACCGCGGGGCCGTCAGCAA 340

Qy 283 AATGACAAACAAGGATGGATCTTGTCTGCAAAAAATGCCAACATTTATGCTGGAATCATC 342

Db 341 GAGGGGAGAGCCGGCTCAACCTGGTCCAGCGAAACGTGAACATCTTCAAGTTTCATCAT 400

Qy 343 CCCAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTGCCACGAATCCTGTCTGAT 402

Db 401 CCCAACATTGTCAAGTACAGTCCACACTGCAAGCTGCTGATGCTCTCCAATCCAGTGGAT 460

Qy 403 GTTTTGACCTATATTAGCTATAAGGCGTCAAGGTTTCCACTAAGCAGAGATTATCGGCTCA 462

Db 461 ATCTGACCTACGTGGCTTGGAAAAATCAGTGGCTTTCCCAAAAACCGAGTAATTGGAAGT 520

Qy 463 GGTACAGTTCTGGATACCTGCTCGTTTAAATACATCCTCGAGAGCACTTCAAGATCTCA 522

Db 521 GGTTGCAATCTGGATTACGCGCGGTTCCGTTACCTGATGGGAGAGAGGCTGGGGTTTCA 580

Qy 523 TCGGACAGCATCGATGCTGTGTAATTGGAGAACATGGTGATTCGGGTGTGCTGTCTGG 582

Db 581 GCGCTGAGCTGTACGCGCTGGTCTCTGGGAGAACATGGCGACTCCAGTGTGCTGTGTGG 640

Qy 583 TCTTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCCAACCATATA 642

Db 641 AGTGGTGTGAATGTTGCCGGCGTCTCCCTGAAGTCTCTTAACCCAGAACTGGGCACTGAC 700

Qy 643 TTTGATCAGAATGCGTTCCATAGAATCTTTGAGCAACCGGAGACGCTGCTTACGATATC 702

Db 701 GCAGACAAGGAGCAGTGAAGAGGTTTCAAGCAGGTTGTTGACAGTGCCTACGAGGTG 760

Qy 703 ATCAAGCGCAAGGCTATATCTTCATATGGAATCG 736

Db 761 ATCAAGCTGAAAGGTTACACATCTCTGGGCCATTG 794

RESULT 7

BI410482

LOCUS

DEFINITION

602963152F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5118756 5', mRNA sequence.

BI410482

VERSION

BI410482.1 GI:15171405

KEYWORDS

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1032)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM11289 row: g column: 13

High quality sequence start: 33

High quality sequence stop: 949.

Location/Qualifiers

FEATURES





Db 376 GAGGGGAGAGCGGCTCAACCTGGTCCAGCGAAACGTGAACATCTTCAAGTTCATCAT 435  
QY 343 CCCAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATGCGCAACATCCTGTGCGAT 402  
Db 436 CCCAACAATTGTCAAGTACAGTCCACACTGCAAGCTGCTGATCGTCTCCAATCCAGTGGAT 495  
QY 403 GTTTTGACCTATATTAGCTATAAGCGCTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCA 462  
Db 496 ATCTTGACCTAGCTGGCTTGGAATAATCACTGGCTTCCCAAAAACCGAGTAATTGGAAGT 555  
QY 463 GGTACAGTTCGTGATACCTGCTCGTTTAAATACATCCTCGGAGGACCTTCAAGATCTCA 522  
Db 556 GGTTCGAATCTGGATTACGCGCGTTCCGTTACCTGATGGGAGAGAGGCTGGGGTTTCAC 615  
QY 523 TCGGACAGCATCGATGCCTGTGTAATTGGAGAACATGGTGATTCGGGTGTCCTGTCTGG 582  
Db 616 GCGCTGAGCTGTCACGGCTGGTCTCGGAGAACATGGCGACTCCAGTGTGCTGTGTGG 675  
QY 583 TCTCTTACCAACATCGACGGCATGAAGTCCGGGATTACTGCGAAAAAGCCAACACATA 642  
Db 676 AGTGGTGAATGTTGCCGGCTCTCCCTGAAGTCTCTTAACCCAGAACTGGGCACCTGAC 735  
QY 643 TTTGATCAGAATGCGTTCCATAGAAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATATC 702  
Db 736 GCAGACAAGGAGCAGTGAAGGAGGTTTCAACAGAGGTGTTGGACAGTGCCTACGAGGTG 795  
QY 703 ATCAAGCGCAAGGCTATATTCTCATATGG 731  
Db 796 ATCAAGCTGAAGGTTACACATCTCTGGG 824

RESULT 9  
LOCUS AU090981 1097 bp mRNA linear EST 23-JAN-2001  
DEFINITION AU090981 lambda ZAPII-Con A stimulated leukocytes Paralicthys olivaceus cDNA clone JFConA516F forward similar to Sphyaena idiaestes Lactate dehydrogenase-A (U80001), mRNA sequence.

ACCESSION AU090981  
VERSION AU090981.1 GI:12391021  
KEYWORDS EST.  
SOURCE bastard halibut.  
ORGANISM Paralicthys olivaceus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes; Pleuronectoidei; Paralicthyidae; Paralicthys.

REFERENCE 1 (bases 1 to 1097)  
AUTHORS Nam,B., Hirono,I. and Aoki,T.  
TITLE Identification of expressed genes from Japanese flounder (Paralicthys olivaceus) leukocytes stimulated with Con A/PMA  
JOURNAL Unpublished (2000)  
COMMENT Contact: Ikuro Hirono  
Laboratory of Genetics and Biochemistry  
Tokyo University of Fisheries  
Konan 4-5-7, Minato-ku, Tokyo 108, Japan  
Email: hirono@tokyo-u-fish.ac.jp.

FEATURES  
source  
1. .1097  
/organism="Paralicthys olivaceus"  
/db\_xref="taxon:8255"  
/clone="JFConA516F"  
/clone\_lib="lambda ZAPII-Con A stimulated leukocytes"  
/cell\_type="leukocytes"  
/dev\_stage="adult"  
/note="common name:japanese flounder ; injected with peptidoglycan"  
BASE COUNT 259 a 293 c 324 g 221 t  
ORIGIN

Query Match 13.5%; Score 131.4; DB 9; Length 1097;  
Best Local Similarity 49.5%; Pred. No. 1.9e-27;  
Matches 373; Conservative 0; Mismatches 371; Indels 9; Gaps 1;

QY 35 CTGATGCCAAAACAGTAAAGTTGTGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAG 94  
Db 121 CTGTTGGCAGCAAGAACCAAGGTACAGTGGTGGCGTGGCATGGTGGCCTCCG 180  
QY 95 CGTATACGTTGCTTCTCAGCGGCATCGTTTCCGAGATTGTCTTATCGACGTGAACAAG 154  
Db 181 CCATGAGCGTCTGCTGAAGGACTTGTGTGATGAGCTGGCCCTGGTTGATGATGGAGG 240  
QY 155 ACAAGCAGAGGGTGAAGCATGGAATTAACCAACCGCAGCACCTTCAAATACAAGGTCTC 214  
Db 241 ACAAGCTGAAGGTGAGGCCATGGACCTGCAGCACGGCTCCCTTCTCAAGACGCACA 300  
QY 215 GAGCGGTGATTATCTCTGACTGCGCTGGCGGGCCA-----TTGTTATTGTCAAT 265  
Db 301 AGATTGTGGCCGACAAAGACTACAGCGTGACAGCCAACCTCAAGGTGGTAGTGGTACTG 360  
QY 266 GTGGATTAAACCAAAAAAATGGAACAAACAAAGGATGGATCTTGTGCAAAAAAATGCCAAC 325  
Db 361 CCGGCGCCCGCAGCAGGAGGGCGAGAGCCGTCTCAACCTGTGTGAGCGTAACGTCAACA 420  
QY 326 TTATGCTGGAATCATCCCCAATGTTGCCAAATATGCTCCTGATACCATCCTGCTTATTG 385  
Db 421 TTTTCAAGTTCATCATCCCCAATGTTCAAGTACAGCCCCCAACTGCATCCTGATGGTGG 480  
QY 386 CCACGAATCCTGTGATGTTTGAACCTATATTAGCTATAAGGCGTCAGGGTTTCCACTAA 445  
Db 481 TCTCAACCCAGTGGACATCTTGACCTACGTGGCCTGGAAGCTGAGCGGCTTCCCCCGTC 540  
QY 446 GCAGAGTTATCGGCTCAGGTAAGTCTGGATACCTGCTGTTTAAATACATCCTCGGAG 505  
Db 541 ACCGCTCATTTGGCTCTGGCAACCAACCTGGACTCTGCCCGTTTCCGCCACCTCATGGGAG 600  
QY 506 AGCACTTCAAGATCTCATCGGACAGCATCGATGCTGTGTAATTGGAGAACATGGTGATT 565  
Db 601 AGAAGCTCCACATCCACCTTCCAGCTGTCTGCTGATGATCGGAGAGCAGGAGACT 660  
QY 566 CGGTGTGCTGTCTGTCTCTTACCAACATCGACGCGCATGAAGCTCCGGGATTACTGCG 625  
Db 661 CCAGTGTGCCCGTGTGGAGCGCGGTGAATGTTGCCGAGTCTTCTGCAAGGTCTGAACC 720  
QY 626 AAAAGCCAAACCATATATTGTGATCAGAATCGGTTCCATAGATCTTTGAGCAACCGGAG 685  
Db 721 CAAAATGGGACTGAGGGTGAAGCGGAGAACTGGAAGCAGGTTCAAGATGGTGGTTG 780  
QY 686 ACGTGTCTTACGATATCATCAAGCGCAAGGCTATATCTTCAATATGAATCGCAGCGGGAT 745  
Db 781 ACGAGCCTATGAGGTATCAAGCTGAAGGGCTACACTTCTGGGCCATTGGCATGTCTG 840  
QY 746 TACTTCGCATAGTAAAGGCGATTTTAGAGGATA 778  
Db 841 TGGCTGACCTGGTGGAAAGCATCTTTGAAGAACA 873

RESULT 10  
LOCUS BQ895574

DEFINITION BQ895574 951 bp mRNA linear EST 16-AUG-2002  
AGENCOURT\_8748093 NIH\_MGC\_130 Mus musculus cDNA clone IMAGE:6335313

5', mRNA sequence.  
ACCESSION BQ895574  
VERSION BQ895574.1 GI:22287588  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 951)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,





Db 417 CCCAACATTTGTCAGTACAGTCCACACTGCAAGCTGCTGATCGTCTCCAATCCAGTGGAT 476

QY 403 GTTTTGGACCTATATTAGCTATAAGCGGTGTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCA 462

Db 477 ATCTTGACCTACGTGGCTTGGAATAATCAGTGGCTTCCCAAAACCGAGTAATTGGAAGT 536

QY 463 GGTACAGTTCTGGATACTGCTGTTTTTAAATACATCCTCGGAGAGCACTTCAAGATCTCA 522

Db 537 GGTGCAATCTGGATTTCAGCGCGGTTCCGTTACCTGATGGGAGAGAGGCTGGGGTTCCAC 596

QY 523 TCGGACAGATCGATCGCTGTGTAATTGGAGAAACATGGTGATTCGGGTGTCCTGTCTGG 582

Db 597 GCGCTGAGCTGTACGGCTGCGGTCTCTGNGAGAAACATGGCGACTCCAGTGTGCTGTGG 656

QY 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCAAACCAATA 642

Db 657 AGTGGTGAATGTTGCCGCGTCTCCCTGAAGTCTCTTAACCCAGAACTGGGCACCTGAC 716

QY 643 TTTGATCAGAATCGTTCCATAGAAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATAATC 702

Db 717 GCAGACAAGGAGCAGTGAAGGAGGTTTCAAGCAGGTTGTGGACAGTGCCTACGAAGTG 776

QY 703 ATCAAGCGCAAGGCTATATCTTCAATGGAATCG 736

Db 777 ATCAAGCTGAAAGGTTACACATCTCTGGGCCATTG 810

RESULT 12

BQ886472

LOCUS

DEFINITION BQ886472 959 bp mRNA linear EST 16-AUG-2002

AGENCY AGENCOURT 8766287 NIH\_MGC\_130 Mus musculus cDNA clone IMAGE:6330399

5', mRNA sequence.

ACCESSION BQ886472

VERSION BQ886472.1 GI:22278486

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 959)

TITLE NIH-MGC <http://mgi.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.

CDNA Library Preparation: ResGen, Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LHAM13784 row: 1 column: 16

High quality sequence stop: 586.

Location/Qualifiers

1..959

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:6330399"

/lab\_host="NIH\_MGC\_130"

/note="Organ: otocysts; Vector: pCMV-SPORT6.1.ccdB; Site\_1: EcorV; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH MGC Library."

FEATURES

source

BASE COUNT 233 a 231 c 268 g 226 t 1 others

ORIGIN

Query Match 13.4%; Score 130; DB 14; Length 959;

Best Local Similarity 50.4%; Pred. No. 4.6e-27;

Matches 350; Conservative 0; Mismatches 335; Indels 9; Gaps 1;

QY 52 AAAGTTGTGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC 111

Db 74 AAGATTACAGTTGTGGGGTTGGTGTCTGTGGCATGGCTTGTGCCATCAGTATCTTAATG 133

QY 112 AGCGGCATCGTTTCCGAGATTGTCTTATCGACGTGAACAAAGACAAAGCAGAGGTTGAA 171

Db 134 AAGGACTTGGCGGATGAGCTTGCCCTTGTGACGTATGGAAGACAAACTCAAGGGCGAG 193

QY 172 AGCATGGACTTAAACCAACGACGACACCTTCAAAATACAAGGTCTCGAGCGGGTATTATCCT 231

Db 194 ATGATGGATCTCCAGCATGGCAGCTCTTCTTAAACACCAAAAATTGTCTCCAGCAA 253

QY 232 GACTGCGTGGCGGGCCA-----TTGTTATTGTACATGTGGGATTAAACCAAAA 282

Db 254 GACTACTGTGTAACACTGCGAACTCCAAGCTGGTCAATTATCACCGCGGGGCCCGTCAGCAA 313

QY 283 AATGGACAAACAAGGATGGATCTTGTCTGCAAAAAATGCCAACATTATGTGGAATCATC 342

Db 314 GAGGGGAGAGCCGGCTCAACCTGGTCCAGCGAAACGTGAACATCTTCAAGTTTCATCAAT 373

QY 343 CCAATGTTCCAAATATGCTCTGTATACCATCTCTGTTATTCACGAATCCTGTCTGAT 402

Db 374 CCAACATTTCAAGTACAGTCCACACTGCAAGCTGTGATCGTCTCCAATCCAGTGGAT 433

QY 403 GTTTTGACCTATATTAGCTATAAGGCGTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCA 462

Db 434 ATCTTGACCTACGTGGCTTGGAAAATCAGTGGCTTTCCAAAAACCGAGTAATTGGAAGT 493

QY 463 GGTACAGTTCTGGATACCTGCTGTTTTTAAATACATCTCGGAGAGCACTTCAAGATCTCA 522

Db 494 GGTGCAATCTGGATTACGCGCGGTTCCGTTACCTGATGGGAGAGAGGCTGGGGTTTCCAC 553

QY 523 TCGGACAGCATCGATGCTGTGTAAATTGGAGAACATGGTGAATTGGGTGTGCTGTGG 582

Db 554 GCGCTGAGCTGTACGCGTGGTCTCTGGGAGAACATGGCGACTCCAGTGTGCTGTGG 613

QY 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTCGAAAAAGCCAAACCATATA 642

Db 614 AGTGGTGTGAATGTTGCCGCGTCTCCCTGAAGTCTCTTAACCCAGAACTGGGCACTGAC 673

QY 643 TTTGATCAGAATCGTTTCCATAGAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATAATC 702

Db 674 GCAGACAAGGAGCAGTGGAGGAGGTTTCAAGCAGGTTGGTGGACAGTGCCTACGAGGTG 733

QY 703 ATCAAGCGCAAGGCTATATCTTCAATGGAATCG 736

Db 734 ATCAAGCCTGAAGGGTACACATCTCTGGGCCATTG 767

RESULT 13

CNS06YKH/c

LOCUS

DEFINITION T3 end of clone AY00AA010D03 of library AY00AA from strain CBS 6340 of Kluyveromyces thermotolerans, genomic survey sequence.

ACCESSION AL4211143

VERSION AL421143.1 GI:12204342

KEYWORDS GSS.

SOURCE Kluyveromyces thermotolerans.

ORGANISM Kluyveromyces thermotolerans

REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Kluyveromyces; Saccharomycetales; Saccharomycetaceae; Kluyveromyces.

AUTHORS 1 (bases 1 to 645)

Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.

TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies

JOURNAL FEBS Lett. 487 (1), 3-12 (2000)

MEDLINE 20584711



PUBMED 11152876  
REFERENCE 2 (bases 1 to 645)  
AUTHORS Malpertuy,A., Llorente,B., Blandin,G., Artiguenave,F., Wincker,P. and Dujon,B.  
TITLE Genomic exploration of the hemiascomycetous yeasts: 10.  
JOURNAL FEBS Lett. 487 (1), 61-65 (2000)  
MEDLINE 20584720  
PUBMED 11152885  
REFERENCE 3 (bases 1 to 645)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces marxianus var. marxianus, Pichia lactis var. lactis, Kluyveromyces thermotolerans, Kluyveromyces angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.  
FEATURES  
source  
1. .645  
/organism="Kluyveromyces thermotolerans"  
/strain="CBS 6340"  
/db\_xref="taxon:4916"  
/clone="AY0AA010D03"  
/clone\_lib="AY0AA"  
/note="end : T3"  
misc\_feature complement(<241. .>471)  
/note="similar to TM1867 [ Thermotoga maritima, L-lactate dehydrogenase ]"  
/evidence=not experimental  
BASE COUNT 179 a 139 c 144 g 181 t 2 others  
ORIGIN  
Query Match 13.4%; Score 129.8; DB 17; Length 645;  
Best Local Similarity 68.0%; Pred. No. 4.4e-27;  
Matches 198; Conservative 0; Mismatches 87; Indels 6; Gaps 1;  
QY 38 ATGCCAAAACAGTAAAGTTGGTGTAGTGGAGTGGGAAGTGTGGGTCTGCCACAGCGT 97  
Db 291 AAGTGAACCCCTGTCAAGGTTGCTGTTATTGTTGTTGGAGTGTGGGCTCTACTACGGCCT 232  
QY 98 ATACGTTGCTTCTCAGCGGCATCGTTTCCGAGATTGTCCTTATCGACGTGAACAAAGACA 157  
Db 231 ACACCTCTGCTATTAGCGAGATGATCTCGGAAGTTGTTGATTGATATCAACACACATA 172  
QY 158 AAGCAGAGGGTGAAGCATGGACTTAAACCCAGCAGCACCTTCAAATACAAGGTCTCGAG 217  
Db 171 AAGCAGAGGGCGGAAGCATGGATTAAACCATGTGTCACCTTCCACAACAGGTTCTGTTG 112  
QY 218 CG-----GGTGATTATCTGACTGCGCTGGCGGCCCATTTGTTATTGTACATGTGGGA 271  
Db 111 TGTATGTTGGAGATTATAGCGACTGCGCCGACGCTGCGGATTGTGATCATACAGGTGGCG 52  
QY 272 TTAACCAAAAAAATGGACAAACAAGGATGGATCTTGTGCAAAAAAATGCCA 322  
Db 51 CCAATCAAAAAACCTGGTCAAACTAGATGGATTAGCAGCAACAAAVGCCA 1  
RESULT 14  
BQ899735  
LOCUS BQ899735  
DEFINITION AGENCOURT\_8752411 NIH\_MGC\_130 Mus musculus cDNA clone IMAGE:6333327 5', mRNA sequence.  
ACCESSION BQ899735  
VERSION BQ899735.1 GI:22291749  
KEYWORDS EST.

SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 926)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.  
cDNA Library Preparation: ResGen, Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLAM13792 row: f column: 16  
High quality sequence stop: 738.  
FEATURES  
source  
1. .926  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6333327"  
/clone\_lib="NIH MGC 130"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: otcysts; Vector: pCMV-SPORT6.1.cdb;  
Site\_1: EcoRV; Site\_2: NotI; Cloned unidirectionally.  
Primer: Oligo dt. Average insert size 1.95 kb.  
Constructed by ResGen, Invitrogen Corp. Note: this is a NIH MGC Library."  
BASE COUNT 218 a 236 c 263 g 209 t  
ORIGIN  
Query Match 13.4%; Score 129.8; DB 14; Length 926;  
Best Local Similarity 50.5%; Pred. No. 5.2e-27;  
Matches 348; Conservative 0; Mismatches 332; Indels 9; Gaps 1;  
QY 52 AAAGTTGTGGTAGTGGGAGTGGGAAGTGTGGGTCTGCCACAGCGTATACGTTGCTTCTC 111  
Db 154 AAGATTACAGTTGTTGGGGTGGTGTGGCATGGCTTGTGCCATCAGTATCTTAATG 213  
QY 112 AGCGGCATCGTTTCCGAGATTGTCCTTATCGACGTGAACAAAGACAAAGCAGAGGGTGAA 171  
Db 214 AAGGACTTGGCGGATGAGCTTGCCCTTGTGACGTATGGAAGACAAACTCAAGGGCGAG 273  
QY 172 AGCATGGACTTAAACCCAGCAGCACCTTCAAATACAAGGTCTCGAGGGGTGATTATCCT 231  
Db 274 ATGATGGATCTCCAGCATGGCAGCCTCTTCTTAAACACCAAAATTTGTCTCCAGCAA 333  
QY 232 GACTCGCTGGCGGCCA-----TTGTTATTGTTCATGTGGGATTAAACCAAAA 282  
Db 334 GACTACTGTGTAACCTGCGAACTCCAAGCTGGTCATTATCACCGGGGGGCCGTCAGCAA 393  
QY 283 AATGGACAAACAAGGATGGATCTGTCTGCAAAAAATGCCAACATTATGCTGGAATCATC 342  
Db 394 GAGGGGAGAGCCGGCTCAACCTGGTCCAGCGAAACGTGAACATCTTCAAGTTCATCAT 453  
QY 343 CCCAATGTTGCCAAATATGCTCTGATACCATCTCTGTTATTGCCACGAATCCTGTGAT 402  
Db 454 CCCAACATTTCAAGTACAGTCCACACTGCAAGCTGCAAGCTGTGATCGTCTCCAATCCAGTGGAT 513  
QY 403 GTTTTGACCTATATTAGCTATAAGGCGTCAAGGTTTCCACTAAGCAGAGATTATCGGCTCA 462  
Db 514 ATCTTGACCTACGTGGCTTGGAAATCAGTGGCTTCCCAAAAAACCGAGTAATTGGAAGT 573  
QY 463 GGTACAGTTCTGGATACCTGCTCGTTTAAATACATCCTCGGAGAGCAGCTTCAAGATCTCA 522  
Db 574 GGTTCGAATCTGGATTTCAGCGCGGTTCCGTTACCTGATGGGAGAGAGGCTGGGGTTTCC 633  
QY 523 TCGGACAGCATCGATGCCTGTGTAATTGGAGAACATGGTGTGATTCGGGTGTGCTGTGG 582

Db 634 GCGCTGAGCTGTACGGCTGGTCTCTGGGGAACATGGCGACTCCAGTGTGCCTGTGTGG 693

QY 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGATTACTGGGAAAAGCCAAACCACATA 642

Db 694 AGTGGTGTGAATGTTGCCGGCTCTCCCTGAAGTCTCTTAACCCAGAACTGGGCACTGAC 753

QY 643 TTTGATCAGATGCGTTCCATAGAATCTTTAGCAAAACGGAGACGCTGCTTACGATATC 702

Db 754 GCAGACAAGGACAGTGAAGGAGGTTTACCAGCAGGTGGTGACAGTGCCTACGAGGTG 813

QY 703 ATCAAGCGCAAGGCTATATCTTCATATGG 731

Db 814 ATCAAGCTGGAAGGTACACATCCTGGG 842

RESULT 15

LOCUS BG916902

DEFINITION 602816178F1 NCI\_CGAP\_Mam6 Mus musculus cDNA clone IMAGE:4935331 5', mRNA sequence.

ACCESSION BG916902

VERSION BG916902.1 GI:14297378

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 924)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

JOURNAL

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgaps@mail.nih.gov](mailto:cgaps@mail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLAM10865 row: 1 column: 20  
High quality sequence stop: 841.

FEATURES

Location/Qualifiers

1..924

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:4935331"

/clone\_lib="NCI\_CGAP Mam6"

/sex="female, virgin"

/tissue\_type="infiltrating ductal carcinoma"

/dev\_stage="5 months"

/lab\_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 231 a 214 c 267 g 212 t

ORIGIN

Query Match 13.3%; Score 129.6; DB 13; Length 924;

Best Local Similarity 50.4%; Pred. No. 6e-27;

Matches 349; Conservative 0; Mismatches 334; Indels 9; Gaps 1;

QY 52 AAAGTTGTGGTGGAGTGGGAAGTGGTCTGCCACAGCGTATACGTTGCTTCTC 111

Db 28 AAGATTACAGTTGTTGGGTTGGTCTGGCATGGCTTGTGCCATCAGTATCTTAATG 87

QY 112 AGCGGCATCGTTCCGAGATTGTCTTATCGACGTGAACAAAGACAAAGCAGGGTGAA 171

Db 88 AAGGACTTGGCGGATGAGCTTGCCCTTGGTTGACGTATGGAAGACAAACTCAAGGGCGAG 147

QY 172 AGCATGGACTTAAACCACGACGACCTTCAAATACAAGTCTCGAGCGGGTATTATCCT 231

Db 148 ATGATGGATCTCCAGCATGGCAGCCTCTTCTTAAACACCAAAAATTGTCTCCAGCAA 207

QY 232 GACTGCGCTGGCGGCCA-----TTGTTATTGTCACTGTGGGATTAAACCAAAA 282

Db 208 GACTACTGTGTAACCTCGAACTCCAAGCTGGTCATTATCACCGCGGGGGCCCGTCAGCAA 267

QY 283 AATGCAAAACAAGGATGGATCTTGTCTGCAAAAAATGCCAACATTATGCTGGAATCATC 342

Db 268 GAGGGGAGAGCCGGCTCAACCTGTCTCCAGCGAAACGTGAACATCTTCAAGTTTCATT 327

QY 343 CCCAATGTTGCCAAATATGCTCTGATACCATCTCTGTTATTGCGACGAATCTGTCTGAT 402

Db 328 CCCAACATTTCAAGTACAGTCCACACTGCAAGCTGCAAGCTGCTCTCTCAATCCAGTGGAT 387

QY 403 GTTTTGACCTATATTAGCTATTAAGCGCTCAGGGTTTCCACTAAGCAGAGTTATCGGCTCA 462

Db 388 ATCTTGACCTACGTGGCTTGGAAAATCAGTGGCTTTCCCAAAAACCGAGTAATTGGAAGT 447

QY 463 GGTACAGTTCTGGATACCTGCTCGTTTAAATACATCTCTCGGAGAGCACTTCAAGATCTCA 522

Db 448 GGTGCAATCTGGATTTCAGCGCGGTTCCGTTACCTGATGGGAGAGAGGCTGGGGTTTCC 507

QY 523 TCGGACAGCATCGATGCCCTGTGTAATTGGAGAACATGGTGATTCGGGTGTGCCTGTCTGG 582

Db 508 GCGCTGAGCTGTACGGCTGGGTCCTGGGAGAACATGGCGACTCCAGTGTGCCTGTGTGG 567

QY 583 TCTCTTACCAACATCGACGGCATGAAGCTCCGGGATTACTGCGAAAAAGCCAAACCACATA 642

Db 568 AGTGGTGTGAATGTTGCCGGCGTCTCCCTGAAGTCTCTTAACCCAGAACTGGGCACTGAC 627

QY 643 TTTGATCAGAATGCGTTCCATAGAATCTTTGAGCAAAACGCGAGACGCTGCTTACGATATC 702

Db 628 GCAGACAAGGAGCAGTGGAAAGAGGTTTCAAAAGCAGGTGGTGACAGTGCCTACGAGGTG 687

QY 703 ATCAAGCGCAAGGCTATATCTTTCATATGGAAT 734

Db 688 ATCAAGCTGAAAGGTTACACATCCTGGGCCAT 719

Search completed: July 31, 2003, 10:16:24

Job time : 1725 secs

